GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 22, 2002, 16:09:32 ; Search time 1.9821 Seconds
(without alignments)
524.976 Million cell updates/sec

Title: US-09-996-738-8
Perfect score: 31
Sequence: 1 VQRGGR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	114	2	C35720	hypothetical 12.4K
2	31	100.0	351	2	C98320	hypothetical prote
3	31	100.0	351	2	AB2963	endoglucanase [imp
4	31	100.0	394	2	I39712	endoglucanase - Ag
5	31	100.0	567	2	I43555	Ras pathway intera
6	31	100.0	1151	2	A45226	integrin alpha 1 c
7	30	96.8	269	2	[illegible]	hypothetical prote
8	30	96.8	1131	2	[illegible]	phytochrome [illegible]
9	28	90.3	134	2	B69156	hypothetical prote
10	28	90.3	235	2	S29652	secretion protein
11	28	90.3	245	2	B83462	probable short-cha
12	28	90.3	281	1	D90151	conserved hypothet
13	28	90.3	283	2	AB1389	B. subtilis YvjA (
14	28	90.3	283	2	AD1764	B. subtilis YvjA (
15	28	90.3	325	1	A48561	inner capsid prote
16	28	90.3	328	1	VPXRC3	minor inner core p
17	28	90.3	338	2	G83349	hypothetical prote
18	28	90.3	375	1	KIBET3	thymidine kinase (
19	28	90.3	393	2	B97128	probable flavoprot
20	28	90.3	449	2	B82958	probable transport
21	28	90.3	481	2	T28990	hypothetical prote
22	28	90.3	502	2	G87433	conserved hypothet
23	28	90.3	509	2	D69832	probable Rieske [2
24	28	90.3	526	2	F29086	hypothetical prote
25	28	90.3	537	2	AG3572	oligopeptide-bindi
26	28	90.3	551	1	HCRVL3	cytochrome c3 hydr
27	28	90.3	577	2	E72296	ABC transporter, A
28	28	90.3	877	2	T03098	p97 protein - Toxo
29	28	90.3	881	1	VCLJG3	env polyprotein -
30	28	90.3	934	2	B88131	protein F10G7.2 [i
31	28	90.3	1729	2	I43103	1,3-beta-glucan sy
32	28	[illegible]	1867	2	[illegible]	probable 1,3-beta
33	27	87.1	31	2	S04980	ferritin heavy cha
34	27	87.1	85	2	S68315	ferritin H chain
35	27	87.1	86	2	B82771	hypothetical prote
36	27	87.1	88	2	I07860	ribosomal protein
37	27	87.1	111	2	[illegible]	adenine phosphorib
38	27	87.1	136	2	F87461	hypothetical prote
39	27	87.1	141	2	I48109	ferritin heavy cha
40	27	87.1	164	2	I46710	ferritin heavy cha
41	27	87.1	165	2	I50725	ribosomal protein
42	27	87.1	174	2	B27805	ferritin chain 2
43	27	87.1	175	1	FRHUL	ferritin light cha
44	27	87.1	175	1	FRHOL	ferritin light cha
45	27	87.1	175	2	S01239	ferritin light cha

ALIGNMENTS

RESULT 1
C35720
hypothetical 12.4K protein - Escherichia coli
N;Alternate names: hypothetical protein 114
C;Species: Escherichia coli
C;Date: [illegible] #sequence_revision [illegible] #text_change 08 Oct 1999
C;Accession: C35720
R;Chen, C.M.; Ye, Q.Z.; Zhu, Z.; Wanner, B.L.; Walsh, C.T.
J. Biol. Chem. 265, 4461-4471, 1990
A;Title: Molecular biology of carbon-phosphorus bond cleavage. Cloning and sequencing
A;Reference number: A35718; MUID:90170953
A;Accession: C35720
A;Molecule type: DNA
A;Residues: 1-114 <CHE>
A;Cross-references: GB:[illegible]; NID:g147[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
C;Genetics:
A;Start codon: GTG
C;Superfamily: Escherichia coli hypothetical 12.4K protein

Query Match 100.0%; Score 31; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
||||||
Db 40 VQRGGR 45

RESULT 2
C98320
hypothetical protein AGR_L_[illegible] [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 22-[illegible]-2002 #sequence_revision 22-[illegible]-2002 #text_change 11-Jan-2002
C;Accession: C98320
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: C98320
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <KUK>
A;Cross-references: GB:AE00[illegible]; PIDN:AAK[illegible]; PID:g[illegible]; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_[illegible]
A;Map position: linear chromosome

Query Match 100.0%; Score 31; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 VQRGGR 6
           ||||||
Db     157 VQRGGR 162

RESULT 3
AB2963
endoglucanase [imported] - Agrobacterium tumefaciens (strain C58, [illegible])
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AB2963
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
[illegible]
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; [illegible]
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB2963
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <KUP
A;Cross-references: GB:AE008689; PIDN:AAL44120.1; PID:g17741690; GSPDB:GN00187
A;Experimental source: strain C58 [illegible]
C;Genetics:
A;Gene: celC
A;Map position: linear chromosome

  Query Match            100.0%;  Score 31;  DB 2;  Length 351;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           ||||||
Db     157 VQRGGR 162

RESULT 4
I39712
endoglucanase - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 08-Oct-1999
C;Accession: I39712
R;Matthysse, A.G.; White, S.; Lightfoot, R.
J. Bacteriol. 177, 1069-1075, 1995
A;Title: Genes required for cellulose synthesis in Agrobacterium tumefaciens.
A;Reference number: I39709; MUID:95164506
A;Accession: I39712
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-394 <RES>
A;Cross-references: GB:L38609; NID:g710486; PIDN:AAC41424.1; PID:g710490
C;Genetics:
A;Gene: celC

  Query Match            100.0%;  Score 31;  DB 2;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 64;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           ||||||
Db     262 VQRGGR 267

RESULT 5
T43555
Ras pathway interacting protein Moe1 - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: elongation initiation factor subunit [misidentification]  Sed1 bindin
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Mar-2000
C;Accession: T43555; T38999
R;Chen, C.R.; Li, Y.C.; Chen, J.; Hou, M.C.; Papadaki, P.; Chang, E.C.
Proc. Natl. Acad. Sci. U.S.A. 96, 517-522, 1999
A;Title: Moe1, a conserved protein in Schizosaccharomyces pombe, interacts with a Ras
A;Reference number: T22574; MUID:99110922
A;Accession: T43555
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-567 <CHE
A;Cross-references: EMBL:AF[illegible]; NID:g[illegible]; PIDN:AAD[illegible]; PID:g[illegible]
R;Purnelle, B.; Goffeau, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21823
A;Accession: T38999
A;Status: [illegible] from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-366,'R',368-567 <PUR>
A;Cross-references: [illegible]
A;Experimental source: strain 972h-; cosmid c637
C;Genetics:
A;Gene: SPDB:SPAC637.07; MOE1
A;Map position: 1
C;Function:
A;Description: a component of the Ras1 pathway necessary for proper spindle formation
A;Note: Moe1 and [illegible] interact directly in vitro [illegible] hypothesized that [illegible]

  Query Match            [illegible];  Score 31;  DB 2;  Length 567;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           ||||||
Db     125 VQRGGR 130

RESULT 6
A45226
integrin alpha 1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 31-[illegible]-2000
C;Accession: A45226
R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A;Title: Expression of native and truncated forms of the human integrin alpha 1 subun
A;Reference number: A45226; MUID:93155124
A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151 <BRI>
A;Experimental source: hepatoblastoma cell line HepG2
A;Note: [illegible] extracted from NCBI backbone (NCBIP:[illegible])
C;Superfamily: [illegible] collagens; von Willebrand factor type A repeat homology
F;142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

  Query Match            [illegible];  Score 31;  DB 2;  Length 1151;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           ||||||
Db     213 VQRGGR 218

RESULT 7
T31318
hypothetical protein - Cenarchaeum symbiosum
C;Species: Cenarchaeum symbiosum
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T31318
R;Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A;Title: Genomic analysis reveals chromosomal variation in natural populations of the
A;Reference number: Z20391; MUID:98422450
A;Accession: T31318
A;Status: preliminary; translated from GB/EMBL/DDBJ
```

A;Molecule type: DNA
A;Residues: 1-269 <SCH>
A;Cross-references: EMBL:AP083772; [illegible]

Query Match 96.8%; Score 30; DB 2; Length 269;
Best Local Similarity 83.3%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 55 IQRGGR 60

RESULT 8
T09701
phytochrome - Scotch pine
C;Species: Pinus sylvestris (Scotch pine)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C;Accession: T09701
R;Wiegmann-Eirund, C.M.; Kolukisaoglu, H.U.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z16826
A;Accession: T09701
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1131 <WIE>
A;Cross-references: EMBL:X96738; NID:g1247883
A;Experimental source: isolate PsA 5.1
C;Superfamily: phytochrome; phytochrome homology
C;Keywords: chromoprotein; photoreceptor; phytochromobilin
F;75-587/Domain: phytochrome homology <PHY>
F;332/Binding site: phytochromobilin [illegible] #status predicted

Query Match 96.8%; Score 30; DB 2; Length 1131;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 82 IQRGGR 87

RESULT 9
B69156
hypothetical protein MTH430 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-[illegible]
C;Accession: B69156
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
[illegible] Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
[illegible] S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514
A;Accession: B69156
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-134 <MTH>
A;Cross-references: GB:AE000827; GB:AE000666; NID:g2621495; PIDN:AAB84937.1; PID:g26214[illegible]
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH430
A;Start codon: TTG

Query Match 90.3%; Score 28; DB 2; Length 134;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 108 IQRGGR 113

RESULT 10
S39652
[illegible] protein xphA [illegible] [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: S39652; H83258
R;Akrim, M.; Bally, M.; Ball, G.; [illegible] A.; Lazdunski,
Mol. Microbiol. 10, 431-443, 1993
A;Title: Xcp-mediated protein secretion in Pseudomonas aeruginosa: identification of
A;Reference number: S39652; MUID:95020542
A;Accession: S39652
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <AKR>
A;Cross-references: EMBL:X68594; NID:g431183; PIDN:CAA48583.1; PID:g431184
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
[illegible]
[illegible] Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: H83258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-[illegible] <STO>
A;Cross-references: GB:AE004[illegible]; [illegible]
A;Experimental source: strain PA01
C;Genetics:
A;Gene: xcpP; PA3104

Query Match 90.3%; Score 28; DB 2; Length 235;
Best Local Similarity 83.3%; Pred. No. 1.[illegible];
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 154 IQRGGR 159

RESULT 11
B83462
probable short chain dehydrogenase PA1470 [imported] - Pseudomonas aeruginosa (strain
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 03-Aug-2001
C;Accession: B83462
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
[illegible]
[illegible] Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: B83462
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-245 <STO>
A;Cross-references: GB:AE[illegible]; GB:AE004[illegible]; NID:g9947[illegible]; PIDN:AAG04859.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA1470
C;Superfamily: [illegible] dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 90.3%; Score 28; DB 2; Length 245;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 129 VERGGR 134

RESULT 12
B30191
conserved hypothetical protein YtIA - Bacillus subtilis

```
C;Species: Bacillus subtilis
C;Date: 15-Oct-1999 #sequence_revision 15-[illegible]-1999 #text_change 21 [illegible]
C;Accession: I40386; D30191; A70042; S28596
R;Londono-Vallejo, J.A.; Dubnau, D.
Mol. Microbiol. 9, 119-131, 1993
A;Title: comF, a Bacillus subtilis late competence locus, encodes a protein similar to [illegible]
A;Reference number: I40386; MUID:94018599
A;Accession: I40386
A;Molecule type: DNA
A;Residues: 1-281 <RES>
A;Cross-references: EMBL:Z18629; NID:g39847; PIDN:CAA79225.1; PID:g39848
R;Henner, D.J.; Yang, M.; Ferrari, E.
J. Bacteriol. 170, 5102-5109, 1988
A;Title: Localization of Bacillus subtilis sacU(Hy) mutations to two linked genes with [illegible]
A;Reference number: A30191; MUID:89033891
A;Accession: D30191
A;Molecule type: DNA
A;Residues: 1-107 <HEN>
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
[illegible]; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
ioch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlan,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
I.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033
A;Accession: A70042
A;Status: nucleic acid sequence not shown, translation not shown
A;Molecule type: DNA
A;Residues: 1-281 <KUN>
A;Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15565.1; PID:g2636074
A;Experimental source: strain 168
C;Genetics:
A;Gene: yviA
C;Superfamily: Mycoplasma hypothetical protein MG326

  Query Match             90.3%;  Score 28;  DB 1;  Length 281;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :|||||
Db        167 LQRGGK 172


RESULT 13
AB1389
B. subtilis YviA (YeqV) protein homolog lmo2[illegible] [imported] - Listeria monocytogenes (st[illegible]
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AB1389
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1389
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00592.1; PID:g16412002; GSPDB:GN00277
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2514

  Query Match             90.3%;  Score 28;  DB 2;  Length 283;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :|||||
Db        169 LQRGGK 174


RESULT 14
AD1764
B. subtilis YviA (YeqV) protein homolog lin2658 [imported] - Listeria innocua (strain
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AD1764
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97884.1; PID:g16415134; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2658

  Query Match             90.3%;  Score 28;  DB 2;  Length 283;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :|||||
Db        169 LQRGGR 174


RESULT 15
A48561
Inner capsid protein VP6 - bluetongue virus (serotype 11, strain USA)
C;Species: bluetongue virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 17-Feb-1994
C;Accession: A48561
R;Hwang, G.Y.; Chiou, J.F.; Yang, Y.Y.; Li, J.K.
Virus Res. 24, 315-323, 1992
A;Title: Comparative sequence analyses of the cognate structural protein VP6 genes of
A;Reference number: A48561; MUID:93033709
A;Accession: A48561
A;Molecule type: genomic RNA
A;Residues: 1-325 <HWA>
A;Note: sequence extracted from NCBI backbone (NCBIN:115442, NCBIP:115444)
C;Genetics:
A;Map position: segment 9
C;Superfamily: bluetongue virus VP6 protein
C;Keywords: capsid protein

  Query Match             90.3%;  Score 28;  DB 1;  Length 325;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :|||||
Db        199 VERGGR 204


Search completed: October 22, 2002, 16:03:24
```

Job Time : 3.09821 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: October 22, 2002, 19:30:01 ; Search Time 0.529429 Seconds
(without alignments)
456.484 Million cell updates/sec

Title: US-09-996-738-8
Perfect score: 41
Sequence: 1 VQRGGR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	100.0	567	1	MOE1_SCHPO	Q94236 schizosacch
2	41	100.0	1151	1	IIA1_HUMAN	P56199 homo sapien
3	40	96.8	1131	1	PHY_PINSY	Q41546 pinus sylv
4	28	90.3	235	1	OSPN_PSEAE	Q51525 pseudomonas
5	28	90.3	281	1	DRGV_BACSU	P32436 bacillus su
6	28	90.3	325	1	VP6_BPVI1	P42933 bluetongue
7	28	90.3	325	1	VP6_BTV10	P32934 bluetongue
8	28	90.3	325	1	VP6_BTV17	P32935 bluetongue
9	28	90.3	329	1	VP61_BTV10	P23066 bluetongue
10	28	90.3	376	1	KITH_HSV23	P04407 herpes simp
11	28	90.3	550	1	PHNI_DESCI	P12944 desulfovibr
12	28	90.3	892	1	ENV_SIVM1	P05885 simian immu
13	27	87.1	88	1	RL15_BRANA	P46289 brassica na
14	27	87.1	153	1	RL15_PIG	P79324 sus scrofa
15	27	87.1	161	1	16SB_RABIT	P25975 oryctolagus
16	27	87.1	165	1	RL15_CHICK	P51417 gallus gall
17	27	87.1	174	1	FRI2_RANCA	P07797 rana catesb
18	27	87.1	174	1	FRIL_BOVIN	Q46415 bos taurus
19	27	87.1	174	1	FRIL_HORSE	P02791 equus cabal
20	27	87.1	174	1	FRIL_HUMAN	P02792 homo sapien
21	27	87.1	174	1	FRIL_RABIT	P09451 oryctolagus
22	27	87.1	177	1	FRIH_SALSA	P69946 salmo salar
23	27	87.1	179	1	FRIH_CHICK	P08267 gallus gall
24	27	87.1	180	1	FRIH_BOVIN	Q46414 bos taurus
25	27	87.1	181	1	FRIH_MOUSE	P09528 mus musculu
26	27	87.1	181	1	FRIH_RAT	P19132 rattus norv
27	27	87.1	182	1	FRIH_HUMAN	P02794 homo sapien
28	27	87.1	182	1	FRIH_TRIVU	Q98173 trichosurus
29	27	87.1	182	1	SRPB_SYNP7	Q55026 synechococc
30	27	87.1	185	1	FRIB_CRIGR	P29389 cricetulus
31	27	87.1	194	1	R15E_PYRAB	Q9V0d2 pyrococcus
32	27	87.1	194	1	R15E_PYRHO	O58706 pyrococcus
33	27	87.1	201	1	RL15_SCHPO	O74895 schizosacch
34	27	87.1	203	1	RL17_ASPNG	Q13418 aspergillus
35	27	87.1	203	1	RL17_CHITE	P48236 chironomus
36	27	87.1	203	1	RL17_HUMAN	P09132 homo sapien
37	27	87.1	204	1	K19A_PEPMA	Q62750 pteromys mi
38	27	87.1	204	1	RT9B_PEPMA	Q63082 pteromys mi
39	27	87.1	204	1	RL17_ARATH	Q39115 arabidopsis
40	27	87.1	204	1	RL17_DROME	P22744 drosophila
41	27	87.1	205	1	RL17_PETMA	Q62738 petromyzon
42	27	87.1	214	1	RL17_MOUSE	Q60605 mus musculu
43	27	87.1	269	1	HI_PEA	P08283 pisum sativ
44	27	87.1	529	1	K6PT_WHEAT	Q40482 triticum ae
45	27	87.1	590	1	DOB1_MYCTU	Q10671 mycobacteri

ALIGNMENTS

RESULT 1
MOE1_SCHPO
ID MOE1_SCHPO STANDARD; PRT; 567 AA.
AC Q94236;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Microtubule destabilizing protein moe1
GN MOE1 OR SPAC664.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99113922; PubMed=9892665;
RA Chen C., Ji Y.-C., Chen L., Bon M.-C., Papadaki P., Chang E.C.;
RT "Moe1, a conserved protein in Schizosaccharomyces pombe, interacts
RT with a Ras effector, Scd1, to affect proper spindle formation.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:517-522(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Purnelle B., Cottrem A., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: POTENTIAL NEGATIVE REGULATOR FOR MICROTUBULE DYNAMICS.
CC -!- PHENOTYPE: INSTABILITY IN MICROTUBULES. AFFECT SPINDLE FORMATION.
CC -!- SUBUNIT: INTERACTS WITH SCD1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE RIB43Z FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF086508; AAD08803.1; -.
DR EMBL; AL031584; CAA22586.1; -.
KW Nuclear protein.
FT NON_TER 1 1 (IN REL. 2)
SQ SEQUENCE 567 AA; 62936 MW; 5C48A2CA754A743B CRC64;

Query Match 100.0%; Score 41; DB 1; Length 567;
Best Local Similarity 100.0%; Pred. No. 3e-
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
|||||| |
Db 127 VQRGGR 132

RESULT 2

```
ITA1_HUMAN
ID   ITA1_HUMAN     STANDARD;      PRT;  1151 AA.
AC   P56199;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN   ITGA1.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93155124; PubMed 8428973.
RA   Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT   "Expression of native and truncated forms of the human integrin alpha
RT   1 subunit.";
RL   J. Biol. Chem. 268:2989-2996(1993).
CC   -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC       COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC       E-R IN COLLAGEN.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC       ASSOCIATES WITH BETA-1.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC       WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC   -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
DR   HSSP; P17301; 1AOX.
DR   MIM; 192968; .
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWFA.
DR   Pfam; PF01839; FG-GAP; 5.
DR   Pfam; PF00357; integrin_A; 1.
DR   Pfam; PF00092; vwa; 1.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW   Repeat; Calcium; Magnesium.
FT   DOMAIN        1   1113       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1114   1136       POTENTIAL.
FT   DOMAIN     1137   1151       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       16     75       FG-GAP 1.
FT   REPEAT [illegible] [illegible]       FG-GAP 2.
FT   DOMAIN      147    360       VWFA.
FT   REPEAT      349    404       FG-GAP 3.
FT   REPEAT      405    457       FG-GAP 4.
FT   REPEAT      459    520       FG-GAP 5.
FT   REPEAT      540    599       FG-GAP 6.
FT   REPEAT      602    654       FG-GAP 7.
FT   CA_BIND     470    478       POTENTIAL.
FT   CA_BIND     552    560       POTENTIAL.
FT   CA_BIND     614    622       POTENTIAL.
FT   SITE       1139   1142       GFFKR MOTIF.
FT   DISULFID     54     64       BY SIMILARITY.
FT   DISULFID    660    669       BY SIMILARITY.
FT   DISULFID    675    728       BY SIMILARITY.
FT   DISULFID    780    786       BY SIMILARITY.
FT   DISULFID    850    858       BY SIMILARITY.
FT   DISULFID   1002   1034       BY SIMILARITY.
FT   DISULFID   1037   1044       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     77     77       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD [illegible] [illegible]       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD [illegible] [illegible]       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    371     71       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD [illegible] [illegible]       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    4[illegible]2    432       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    5[illegible]4    5[illegible]4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    571    571       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    720    720       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    752    752       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    8[illegible]2    8[illegible]2       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    855    855       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    880    880       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    887    887       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    911    911       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD [illegible] [illegible]       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    946    946       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD [illegible]    [illegible]       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1045   1045       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1055   1055       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1071   1071       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1[illegible]   1085       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1151 AA;  127837 MW;  6B3F3C1AABF52808 CRC64;

  Query Match             100.0%;  Score 31;  DB 1;  Length 1151;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||||||
Db        214 VQRGGR 218


RESULT 3
PHY_PINSY
ID   PHY_PINSY      STANDARD;      PRT;  1141 AA.
AC   Q41046;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Phytochrome.
OS   Pinus sylvestris (Scots pine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=3349;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PSA 5.1;
RA   Wiegmann-Eirund C.M., Kolukisaoglu H.U.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
CC       ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
CC       MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
CC       ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR TO
CC       PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
CC       RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE
CC       RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
CC       GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-
CC       BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
CC       PROTOCHLOROPHYLLIDE REDUCTASE, RRNA, ETC. IT ALSO CONTROLS
CC       THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
CC   -!- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 PAS (PER ARNT SIM) DIMERIZATION DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 PAS ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   [illegible]
CC   [illegible] requires a license agreement [illegible]
```

```
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X96738; CAA65510.1; -.
DR   InterPro; IPR003018; GAF.
DR   InterPro; IPR003594; HATPase_c.
DR   InterPro; IPR004359; HIS_KIN_sig.
DR   InterPro; IPR003661; His_kinA.
DR   InterPro; IPR001610; PAC.
DR   InterPro; IPR000014; PAS.
DR   InterPro; IPR001294; Phytochrome.
DR   Pfam; PF01590; GAF; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00989; PAS; 2.
DR   Pfam; PF00360; phytochrome; 1.
DR   Pfam; PF00512; signal; 1.
DR   PRINTS; PR01033; PHYTOCHROME.
DR   SMART; SM00065; GAF; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00086; PAC; 1.
DR   SMART; SM00091; PAS; 2.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50112; PAS; 2.
DR   PROSITE; PS00245; PHYTOCHROME_1; 1.
DR   PROSITE; PS50046; PHYTOCHROME_2; 1.
KW   Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW   Repeat.
FT   DOMAIN      621    692       PAS 1.
FT   DOMAIN      755    826       PAS 2.
FT   DOMAIN      903   1123       HISTIDINE KINASE.
FT   BINDING     332    332       CHROMOPHORE (BY SIMILARITY).
SQ   SEQUENCE   1131 AA;  126254 MW;  B63A2 8FA9862FB CRC64;

  Query Match             96.8%;  Score 30;  DB 1;  Length 1131;
  Best Local Similarity   83.3%;  Pred. No. 80;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db         82 IQRGGR 87


RESULT 4
GSPN_PSEAE
ID   GSPN_PSEAE     STANDARD;      PRT;   235 AA.
AC   Q51575; Q9HZI3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   General secretion pathway protein N (XcpP protein).
GN   XCPP OR PA3104.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=95020512; PubMed=7934833;
RA   Akrim M., Bally M., Ball G., Tommassen J., Teerink H., Filloux A.,
RA   Lazdunski A.;
RT   "Xcp-mediated protein secretion in Pseudomonas aeruginosa:
RT   identification of two additional genes and evidence for regulation of
RT   xcp gene expression.";
RL   Mol. Microbiol. 10:431-443(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
RN   [3]
RP   SUBCELLULAR LOCATION, AND TOPOLOGY.
RX   MEDLINE=96412375; PubMed=8763961;
RA   Bleves S., Lazdunski A., Filloux A.;
RT   "Membrane topology of three Xcp proteins involved in exoprotein
RT   transport by Pseudomonas aeruginosa.";
RL   J. Bacteriol. 178:4297-4300(1996).
CC   -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC       EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF A VARIETY
CC       OF ENZYMES ACROSS THE OUTER MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X68594; CAA48581.1; -.
DR   EMBL; AE004729; AAG06492.1; -.
KW   Inner membrane; Transmembrane; Transport; Signal-anchor;
KW   Complete proteome.
FT   DOMAIN        1     34       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     35     55       SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       56    235       PERIPLASMIC (POTENTIAL).
SQ   SEQUENCE   235 AA;  25496 MW;  DE9726984D89A0D7 CRC64;

  Query Match             93.3%;  Score 28;  DB 1;  Length 235;
  Best Local Similarity   83.3%;  Pred. No. 1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        154 IQRGGR 159


RESULT 5
BEGV_BACSU
ID   BEGV_BACSU     STANDARD;      PRT;   281 AA.
AC   P39146;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   BEGV protein.
GN   BEGV
OS   Bacillus subtilis
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=94185909; PubMed=8412657;
RA   Londono-Vallejo J.A., Dubnau D.;
RT   "comF, a Bacillus subtilis late competence locus, encodes a protein
RT   similar to ATP-dependent RNA/DNA helicases.";
RL   Mol. Microbiol. 9:119-131(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Sorokin B., Lazarevic V., Mauel C., Karamata D.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-137 FROM N.A.
RX   MEDLINE=20403051; PubMed=8141378;
```

```
RA   Henner D.J., Yang M., Ferrari E.;
RT   "Localization of Bacillus subtilis sacU(Hy) mutations to two linked
RT   genes with similarities to the conserved procaryotic family of two
RT   component signalling systems.";
RL   J. Bacteriol. 170:5102-5109(1988).
CC   -!- SIMILARITY: TO F AURPPS PLASMID PFI HYPOTHETICAL 15.5 KD PROTEIN
CC       (AC P13977) AND TO M.GENITALIUM MG126.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z18629; CAA79225.1; -.
DR   EMBL; U56901; AAC44939.1; -.
DR   EMBL; M23558; -; NOT_ANNOTATED_CDS.
DR   EMBL; Z99122; CAB15565.1; -.
DR   PIR; D30191; D30191.
DR   PIR; S28596; S28596.
DR   SubtiList; BG10394; degV.
DR   InterPro; IPR003797; DUF194.
DR   Pfam; PF02645; DUF194; 1.
KW   Complete proteome.
SQ   SEQUENCE   281 AA;  31590 MW;  [illegible] CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 281;
  Best Local Similarity   83.3%;  Pred. No. 59;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        167 LQRGGR 172

RESULT 6
VP6_BTV11
ID   VP6_BTV11      STANDARD;      PRT;   325 AA.
AC   P32933;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VP6 protein (Minor inner core protein VP6).
GN   S9.
OS   Bluetongue virus (serotype 11 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=33716;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93033709; PubMed=1329371;
RA   Hwang G.-Y., Chiou J.-F., Yang Y.-Y., Li J.K.-K.;
RT   "Comparative sequence analyses of the cognate structural protein VP6
RT   genes of five US bluetongue viruses.";
RL   Virus Res. 24:315-323(1992).
CC   -!- FUNCTION: SURROUNDS AND INTERACTS WITH THE GENOMIC DS RNA.
CC       POSSESSES SS- AND DS-RNA-BINDING CAPACITY. ITS HYDROPHILIC
CC       NATURE AND CAPABILITY TO BIND SS- AND DS-RNA SUGGEST ITS
CC       INTERACTIONS WITH BTV GENOMIC RNA.
CC   -!- SUBCELLULAR LOCATION: INNER CAPSID.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08670; AAA42819.1; -.
DR   PIR; A48561; A48561.
DR   InterPro; IPR001399; Orbi_VP6.
DR   Pfam; PF01516; Orbi_VP6; 1.
DR   PRINTS; PR00902; VP6CAPSID.
KW   Core protein.
FT   DOMAIN       84    125       GLY RICH.
SQ   SEQUENCE   325 AA;  35271 MW;  [illegible] CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 325;
  Best Local Similarity   83.3%;  Pred. No. 67;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        199 VERGGR 204

RESULT 7
VP6_BTV13
ID   VP6_BTV13      STANDARD;      PRT;   325 AA.
AC   P32934;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VP6 protein (Minor inner core protein VP6).
GN   S9.
OS   Bluetongue virus (serotype 13 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=33717;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93033709; PubMed=1329371;
RA   Hwang G.-Y., Chiou J.-F., Yang Y.-Y., Li J.K.-K.;
RT   "Comparative sequence analyses of the cognate structural protein VP6
RT   genes of five US bluetongue viruses.";
RL   Virus Res. 24:315-323(1992).
CC   -!- FUNCTION: SURROUNDS AND INTERACTS WITH THE GENOMIC DS RNA.
CC       POSSESSES SS- AND DS-RNA-BINDING CAPACITY. ITS HYDROPHILIC
CC       NATURE AND CAPABILITY TO BIND SS- AND DS-RNA SUGGEST ITS
CC       INTERACTIONS WITH BTV GENOMIC RNA.
CC   -!- SUBCELLULAR LOCATION: INNER CAPSID.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08671; AAA42820.1; -.
DR   InterPro; IPR001399; Orbi_VP6.
DR   Pfam; PF01516; Orbi_VP6; 1.
DR   PRINTS; PR00902; VP6CAPSID.
KW   Core protein.
FT   DOMAIN       84    125       GLY RICH.
SQ   SEQUENCE   325 AA;  35245 MW;  P0F4DB7127[illegible]A334 CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 325;
  Best Local Similarity   83.3%;  Pred. No. 67;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        199 VERGGR 204

RESULT 8
VP6_BTV17
ID   VP6_BTV17      STANDARD;      PRT;   325 AA.
AC   P32935;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
```

```
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VP6 protein (Minor inner core protein VP6).
GN   S9.
OS   Bluetongue virus (serotype 17 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID 33718;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 93033709; PubMed 1329371;
RA   Hwang G. Y., Chiou J. F., Yang Y. Y., Li J.K. K.;
RT   "Comparative sequence analyses of the cognate structural protein VP6
RT   genes of five US bluetongue viruses.";
RL   Virus Res. 24:315-323(1992).
CC   -!- FUNCTION: SURROUNDS AND INTERACTS WITH THE GENOMIC DS RNA.
CC       POSSESSES SS- AND DS-RNA BINDING CAPACITY. ITS HYDROPHILIC
CC       NATURE AND CAPABILITY TO BIND SS- AND DS-RNA SUGGEST ITS
CC       INTERACTIONS WITH BTV GENOMIC RNA.
CC   -!- SUBCELLULAR LOCATION: INNER CAPSID.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08672; AAA42821.1; -.
DR   InterPro; IPR001399; Orbi_VP6.
DR   Pfam; PF01516; Orbi_VP6; 1.
DR   PRINTS; PR00902; VP6CAPSID.
KW   Core protein.
FT   DOMAIN       84    125       GLY RICH.
SQ   SEQUENCE   325 AA;  35317 MW;  6E8C5985A2DA11C1 CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 325;
  Best Local Similarity   83.3%;  Pred. No. 67;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                |:||||
Db        199 VERGGR 204


RESULT 9
VP61_BTV10
ID   VP61_BTV10     STANDARD;      PRT;   329 AA.
AC   P23066;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VP6 protein (Minor inner core protein VP6) (Version 1).
GN   S9.
OS   Bluetongue virus (serotype 10 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID 10900;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 89293876; PubMed 2544665;
RA   Fukusho A., Yu Y., Yamaguchi S., Roy P.;
RT   "Completion of the sequence of bluetongue virus serotype 10 by the
RT   characterization of a structural protein, VP6, and a non structural
RT   protein, NS2.";
RL   J. Gen. Virol. 70:1677-1689(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 90349726; PubMed 2166648;
RA   Roy P., Marshall J.J.A., French T.J.;
RT   "Structure of the bluetongue virus genome and its encoded proteins.";
RL   Curr. Top. Microbiol. Immunol. 162:43-87(1990).
RN   [3]
RP   REVISIONS
RA   Bluchem ...
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SURROUNDS AND INTERACTS WITH THE GENOMIC DS RNA.
CC       POSSESSES SS- AND DS-RNA BINDING CAPACITY. ITS HYDROPHILIC
CC       NATURE AND CAPABILITY TO BIND SS- AND DS-RNA SUGGEST ITS
CC       INTERACTIONS WITH BTV GENOMIC RNA.
CC   -!- SUBCELLULAR LOCATION: INNER CAPSID.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC   -!- CAUTION: SEE ALSO VERSION 2 OF THIS PROTEIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D00509; BAA00399.1; -.
DR   PIR; B32400; VPXRC3.
DR   PIR; S10541; S10542.
DR   InterPro; IPR001399; Orbi_VP6.
DR   Pfam; PF01516; Orbi_VP6; 1.
DR   PRINTS; PR00902; VP6CAPSID.
KW   Core protein.
FT   DOMAIN       88    130       GLY RICH.
SQ   SEQUENCE   329 AA;  35515 MW;  5BDA9F44D7E2343D CRC64;

  Query Match             90.3%;  Score 28;  DB 1;  Length 329;
  Best Local Similarity   83.3%;  Pred. No. 68;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                |:||||
Db        203 VERGGR 208


RESULT 10
KITH_HSV23
ID   KITH_HSV2      STANDARD;      PRT;   376 AA.
AC   P04407;
DT   13-AUG-1987 (Rel. 05, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Thymidine kinase (EC 2.7.1.21).
GN   TK.
OS   Herpes simplex virus (type 2 / strain 333).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID 10313;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 84080453; PubMed 6317035;
RA   Kit S., Kit M., Qavi H., Trkula D., Otsuka H.;
RT   "Nucleotide sequence of the herpes simplex virus type 2 (HSV-2)
RT   thymidine kinase gene and predicted amino acid sequence of thymidine
RT   kinase polypeptide and its comparison with the HSV-1 thymidine kinase
RT   gene.";
RL   Biochim. Biophys. Acta 741:158-170(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 83216300; PubMed 6304336;
RA   Swain M.A., Galloway D.A.;
RT   "Nucleotide sequence of the herpes simplex virus type 2 thymidine
RT   kinase gene.";
RL   J. Virol. 46:1045-1050(1983).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 88132670; PubMed 2829705.
RA   Kit S., Sheppard M., Ichimura H., Nishioka Y., Feldman C., Flint M.N.,
RA   Lyle J.A., Otsuka H.;
RT   "Nucleotide sequence changes in thymidine kinase gene of herpes
RT   simplex virus type 2 clones from an isolate of a patient treated with
```

```
RT   acyclovir.";
RL   Antimicrob. Agents Chemother. 31:1483-1490(1987).
CC   -!- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-
CC       phosphate.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRIDAE THYMIDINE KINASE FAMILY
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61713; CAA43858.1; -.
DR   EMBL; V00466; CAA23749.1; -.
DR   EMBL; M29942; AAA45856.1; -.
DR   PIR; A00613; KIBE13.
DR   HSSP; P03176; 2VTK.
DR   InterPro; IPR000980; TK_herpes.
DR   Pfam; PF00693; TK_herpes; 1.
DR   ProDom; PD001519; TK_herpes; 1.
KW   Transferase; Kinase; DNA synthesis; ATP binding.
FT   NP_BIND      56     63       ATP (PROBABLE).
FT   DOMAIN      168    176       NUCLEOSIDE-BINDING (POTENTIAL).
FT   CONFLICT    271    271       MISSING (IN REF. 1).
SQ   SEQUENCE   376 AA;  40467 MW;  0B1D68B0BBA398A5 CRC64;

  Query Match             99.8%;  Score 28;  DB 1;  Length 376;
  Best Local Similarity   83.3%;  Pred. No. 77;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 VQRGGR 6
           :||||:
Db     250 LQRGGR 255


RESULT 11
PHNL_DESGI
ID   PHNL_DESGI     STANDARD;      PRT;   550 AA.
AC   P12944;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Periplasmic [NiFe] hydrogenase large subunit precursor (EC 1.18.99.1)
DE   (NiFe hydrogenase large chain).
GN   HYDB.
OS   Desulfovibrio gigas.
OC   Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
OX   NCBI_TaxID=879;
RN   [1]
RP   SEQUENCE FROM N.A.  AND REVISIONS.
RX   MEDLINE; 88111029; PubMed=3122743;
RA   Li C., Peck H.D. Jr., Le Gall J., Przybyla A.E.;
RT   "Cloning, characterization, and sequencing of the genes encoding the
RT   large and small subunits of the periplasmic [NiFe]hydrogenase of
RT   Desulfovibrio gigas.";
RL   DNA 6:539-551(1987).
RN   [2]
RP   SEQUENCE FROM N.A., AND REVISIONS.
RX   MEDLINE; 89213580; PubMed=2521412;
RA   Voordouw G., Menon N.K., Le Gall J., Choi E.S., Peck H.D. Jr.,
RA   Przybyla A.E.;
RT   "Analysis and comparison of nucleotide sequences encoding the genes
RT   for [NiFe] and [NiFeSe] hydrogenases from Desulfovibrio gigas and
RT   Desulfovibrio baculatus.";
RL   J. Bacteriol. 171:2894-2899(1989).
RN   [3]
RP   SEQUENCE OF 1-29.
RX   MEDLINE; 88106446; PubMed=3222275.
RA   Prickril B.C., Ro S.H., Li C., Menon N., Choi E.S., Przybyla A.E.;
RA   Dervartanian D.V., Peck H.D. Jr., Huynh B.H., Le Gall J., Voordouw G.,
RA   Moura I., Moura J.J.G., Patil D., Rajagopalan K.V.;
RT   "Identification of three classes of hydrogenase in the genus
RT   Desulfovibrio.";
RL   Biochem. Biophys. Res. Commun. 149:369-377(1987).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX   MEDLINE=95157629; PubMed=7854413;
RA   Volbeda A., Charon M.-H., Piras C., Hatchikian E.C., Frey M.,
RA   Fontecilla-Camps J.C.;
RT   "Crystal structure of the nickel-iron hydrogenase from Desulfovibrio
RT   gigas.";
RL   Nature 373:580-587(1995).
CC   -!- CATALYTIC ACTIVITY: H(2) + 2 ferricytochrome b(-557) = 2 H(+) + 2
CC       ferrocytochrome b(-557).
CC   -!- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE 3FE-4S
CC       CLUSTER. THE LARGE SUBUNIT IS INVOLVED IN MAIN THE NICKEL ION.
CC   -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- MISCELLANEOUS: PERHAPS THE LEADER OF THE SMALL SUBUNIT SERVES AS A
CC       TRANSPORT VEHICLE FOR BOTH SUBUNITS.
CC   -!- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE
CC       SUBUNIT FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M1-83; AAA23478.1; ALT_SEQ.
DR   PIR; B32415; B32415.
DR   PIR; D27480; D27480.
DR   PDB; 1FRV; 08-NOV-96.
DR   InterPro; IPR001501; NiFeSe_Hases.
DR   Pfam; PF00374; NiFeSe_Hases; 1.
DR   PROSITE; PS00507; NI_HGENASE_L_1; 1.
DR   PROSITE; PS00508; NI_HGENASE_L_2; 1.
KW   Oxidoreductase; Periplasmic; Nickel; 3D structure.
FT   INIT_MET      0      0
FT   CHAIN         1    535       PERIPLASMIC [NIFE] HYDROGENASE LARGE
FT                                SUBUNIT.
FT   PROPEP      536    550
FT   METAL        64     64       NICKEL.
FT   METAL        67     67       NICKEL.
FT   METAL       529    529       NICKEL.
FT   METAL       532    532       NICKEL.
SQ   SEQUENCE   550 AA;  61449 MW;  [illegible] CRC64;

  Query Match             99.8%;  Score 28;  DB 1;  Length 550;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 VQRGGR 6
           :||||:
Db     470 VQRGGK 475


RESULT 12
ENV_SIVM1
ID   ENV_SIVM1      STANDARD;      PRT;   856 AA.
AC   P05895; Q85725; Q85726;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Simian immunodeficiency virus (Mm142-83 isolate) (SIV-MAC).
OC   Viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11733;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE 87287230; PubMed 3649526;
RA   Chakrabarti L., Guyader M., Alizon M., Daniel M.D., Desrosiers R.C.,
RA   Tiollais P., Sonigo P.;
RT   "Sequence of simian immunodeficiency virus from macaque and its
RT   relationship to other human and simian retroviruses";
RL   Nature 328:543-547(1987).
CC   -!- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Y00277; . NOT_ANNOTATED_CDS.
DR   PIR; B28887; VCLJG3.
DR   HIV; MI6403; ENVSMM142.
DR   InterPro; IPR000328; Env_GP41.
DR   Interfil; IPR003772; S12.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    528       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       529    802       TRANSMEMBRANE GLYCOPROTEIN.
FT   SITE        747    747       IN FRAME TERMINATION CODON.
FT   CARBOHYD     37     37       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     70     70       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    149    149       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    159    159       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    174    174       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    201    201       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    205    205       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    215    215       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    281    281       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    287    287       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    298    298       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    319    319       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    374    374       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    380    380       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    463    463       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    474    474       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    479    479       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    628    628       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    637    637       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    653    653       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    762    762       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   882 AA;  101196 MW;  2160D5206C805DA7 CRC64;

  Query Match             97.8%;  Score 2%;  DB 1;  Length 882;
  Best Local Similarity   86.3%;  Pred. No. 1.2e-02;
  Matches   5;  Conservative   1;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1 VQRGGR 6
           :|||||
Db     859 LQRGGR 864


RESULT 13
RL15_BRANA
ID   RL15_BRANA     STANDARD;      PRT,    88 AA.
AC   P46289;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   [illegible] (Rel. 33, Last annotation update)
DE   60S ribosomal protein L15 (RL15) (Fragment)
GN   RPL15.
OS   Brassica napus (Rape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SAMOURAI;
RA   Saez-Vasquez J.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RL   [illegible]
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.[illegible]/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U21246; AAA86368.1; -.
DR   InterPro; IPR000439; Ribosomal_L15e.
DR   Pfam; PF00827; Ribosomal_L15e; 1.
DR   PROSITE; PS01194; RIBOSOMAL_L15E; 1.
KW   Ribosomal protein.
FT   NON_TER       1      1
FT   NON_TER      88     88
SQ   SEQUENCE   88 AA;  10445 MW;  0F3443489790B0D0 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 88;
  Best Local Similarity   86.3%;  Pred. No. [illegible];
  Matches   5;  Conservative   1;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1 VQRGGR 6
           :|||||
Db      [illegible] VKKGGR [illegible]


RESULT 14
RL15_BRANA
ID   RL15_BRANA     STANDARD;      PRT,    153 AA.
AC   P46624;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   60S ribosomal protein L15 (Fragment).
GN   RPL15.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Small intestine;
RA   Winterroe A.K., Brodholm M.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RL   -!- SIMILARITY: BELONGS TO THE L15E FAMILY OF RIBOSOMAL PROTEINS.
CC
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.[illegible]/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Z84172; CAB06330.1; -.
DR   InterPro; IPR000439; Ribosomal_L15e.
DR   Pfam; PF00827; Ribosomal_L15e; 1.
DR   PROSITE; PS01194; RIBOSOMAL_L15E; PARTIAL.
```

```
KW   Ribosomal protein.
FT   NON_TER       1       1
SQ   SEQUENCE   153 AA;  [illegible] MW;  A[illegible] CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 153;
  Best Local Similarity   83.3%;  Pred. No. 55;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           |:|||| 
Db      15 VRRGGR 20


RESULT 15
FRIH_RABIT
ID   FRIH_RABIT     STANDARD;      PRT;   164 AA.
AC   P25915;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ferritin heavy chain (Ferritin H subunit) (Fragment)
GN   FTH.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Vascular smooth muscle;
RX   MEDLINE=92011647; PubMed=1655764;
RA   Liau G., Chan L.M., Feng P.;
RT   "Increased ferritin gene expression is both promoted by cAMP and a
RT   marker of growth arrest in rabbit vascular smooth muscle cells.";
RL   J. Biol. Chem. 266:18819-18826(1991).
CC   -!- FUNCTION: FERRITIN IS AN INTRACELLULAR PROTEIN THAT STORES IRON
CC       IN A SOLUBLE, NONTOXIC, READILY AVAILABLE FORM. THE FUNCTIONAL
CC       MOLECULE, WHICH IS COMPOSED OF 24 CHAINS, IS ROUGHLY SPHERICAL
CC       AND CONTAINS A CENTRAL CAVITY IN WHICH THE POLYMERIC FERRIC IRON
CC       CORE IS DEPOSITED.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF FERRITIN SUBUNITS: L (LIGHT)
CC       CHAIN AND H (HEAVY) CHAIN. THE MAJOR CHAIN CAN BE LIGHT OR HEAVY,
CC       DEPENDING ON THE SPECIES AND TISSUE TYPE.
CC   -!- SIMILARITY: BELONGS TO THE FERRITIN FAMILY.
CC   -  ---------------------- ---------------------- ---------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaborat on
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   ---------------------------------- ---- --- -- --- ----------
DR   EMBL; M63912; AAA31247.1; -.
DR   HSSP; P02794; 2FHA.
DR   InterPro; IPR001519; Ferritin.
DR   Pfam; PF00210; ferritin; 1.
DR   ProDom; PD000971; Ferritin; 1.
DR   PROSITE; PS00540; FERRITIN_1; 1.
DR   PROSITE; PS00204; FERRITIN_2; 1.
KW   Iron storage; Multigene family.
FT   NON_TER       1       1
FT   METAL         9       9       IRON (BY SIMILARITY).
FT   METAL        40      40       IRON (BY SIMILARITY).
FT   METAL        43      43       IRON (BY SIMILARITY).
FT   METAL        44      44       IRON (BY SIMILARITY).
FT   METAL        46      46       IRON (BY SIMILARITY).
FT   METAL        47      47       IRON (BY SIMILARITY).
FT   METAL        89      89       IRON (BY SIMILARITY).
FT   METAL       123     123       IRON (BY SIMILARITY).
SQ   SEQUENCE   164 AA;  19192 MW;  C1805C5CE8FBB389 CRC64;

  Query Match             87.1%;  Score 27;  DB 1;  Length 164;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRGGR 6
           |||||
Db      57 QRGGR 61


Search completed: October 22, 2002, 16:00:35
Job time : 1.50894 secs
```

GeneCore version 5.1.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2002, 16:00:02 ; Search time 1.63393 Seconds
(without alignments)
635.260 Million cell updates/sec

Title: US-09-996-738-8
Perfect score: 31
Sequence: 1 VQRGGR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994629 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrates:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	96	15	Q91B97	Q91B97 human immun
2	31	100.0	114	2	Q47482	Q47482 escherichia
3	31	100.0	278	2	Q07397	Q07397 mycobacteri
4	31	100.0	394	2	Q44415	Q44415 agrobacteri
5	31	100.0	424	16	Q98BU6	Q98BU6 rhizobium l
6	30	96.8	130	1	Q9UWU6	Q9UWU6 sulfolobus
7	30	96.8	269	1	O74066	O74066 cenarchaeum
8	30	96.8	273	1	O74043	O74043 cenarchaeum
9	30	96.8	273	2	Q9AEN0	Q9AEN0 streptomyce
10	28	90.3	72	2	Q93N80	Q93N80 streptomyce
11	28	90.3	134	17	O26530	O26530 methanother
12	28	90.3	141	4	Q9RX98	Q9RX98 homo sapien
13	28	90.3	212	16	Q94IP0	Q94IP0 [illegible]
14	28	90.3	215	16	Q93HP2	Q93HP2 pseudomonas
15	28	90.3	283	16	Q927X9	Q927X9 [illegible]
16	28	90.3	299	3	Q9HGE5	Q9HGE5 schizosacch
17	28	90.3	326	1	Q9HHI6	Q9HHI6 methanosarc
18	28	90.3	329	12	Q98816	Q98816 bluetongue
19	28	90.3	329	12	Q98824	Q98824 bluetongue
20	28	90.3	329	12	Q98829	Q98829 bluetongue
21	28	90.3	329	12	Q98836	Q98836 bluetongue
22	28	90.3	329	12	Q96648	Q96648 bluetongue
23	28	90.3	329	12	Q96649	Q96649 bluetongue
24	28	90.3	329	12	Q96650	Q96650 bluetongue
25	28	90.3	329	12	Q96651	Q96651 bluetongue
26	28	90.3	329	12	Q96652	Q96652 bluetongue
27	28	90.3	329	12	Q96653	Q96653 bluetongue
28	28	90.3	329	12	Q96654	Q96654 bluetongue
29	28	90.3	329	12	Q96655	Q96655 bluetongue
30	28	90.3	329	12	Q96656	Q96656 bluetongue
31	28	90.3	329	12	Q96657	Q96657 bluetongue
32	28	90.3	329	12	Q96658	Q96658 bluetongue
33	28	90.3	329	12	Q96661	Q96661 bluetongue
34	28	90.3	329	12	Q96664	Q96664 bluetongue
35	28	90.3	368	16	Q91IA9	Q91IA9 pseudomonas
36	28	90.3	375	12	Q9Q0E1	Q9Q0E1 herpes simp
37	28	90.3	375	12	Q9Q0L0	Q9Q0L0 herpes simp
38	28	90.3	376	12	Q9Q0E9	Q9Q0E9 herpes simp
39	28	90.3	376	12	Q69110	Q69110 herpes simp
40	28	90.3	376	12	P89446	P89446 herpes simp
41	28	90.3	376	12	Q91HH5	Q91HH5 herpes simp
42	28	90.3	393	16	Q97107	Q97107 clostridium
43	28	90.3	449	16	Q9HI63	Q9HI63 pseudomonas
44	28	90.3	479	5	Q22549	Q22549 caenorhabdi
45	28	90.3	502	16	Q9A876	Q9A876 caulobacter

ALIGNMENTS

RESULT 1
Q91B97
ID Q91B97 PRELIMINARY; PRT; 96 AA.
AC Q91B97;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VER PROTEIN
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses; Retro-ll viruses; Retroviridae; Lentivirus
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ON340_5
RA Moriyama H.;
RT "HIV-1 in Japan.";
RL Submitted (MAR-1997) to EMBL/GenBank/DDBJ databases.
DR EMBL; AB061186; BAA61996.1; -.
DR Interpro; IPR000012; HIV_VPRVPX.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVVPRVPX.
SQ SEQUENCE 96 AA; 11265 MW; 43FBFA5A7BE0841F CRC64;

Query Match 100.0%; Score 31; DB 15; Length 96;
Best Local Similarity 100.0%; Pred. No. 19
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
||||||
Db 91 VQRGGR 96

RESULT 2
Q47482
ID Q47482 PRELIMINARY; PRT; 114 AA.
AC Q47482;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORF114.
OS   Escherichia coli
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90170953; PubMed=2155230;
RA   Chen C.-M., Ye Q.-Z., Zhu Z., Wanner B.L., Walsh C.T.;
RT   "Molecular biology of carbon-phosphorus bond cleavage. Cloning and
RT   sequencing of the phn (psiD) genes involved in alkylphosphonate uptake
RT   and C-P lyase activity in Escherichia coli B.";
RL   J. Biol. Chem. 265:4461-4471(1990).
DR   EMBL; J05260; AAA24351.1; -.
SQ   SEQUENCE   114 AA;  12496 MW;  F754EC4261530I1A CRC64;

  Query Match             100.0%;  Score 31;  DB 2;  Length 114;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||||||
Db         40 VQRGGR 45

RESULT 3
O07397
ID   O07397      PRELIMINARY;      PRT;   278 AA.
AC   O07397;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MAV278.
GN   MAV278.
OS   Mycobacterium avium.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1764;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GIR10;
RX   MEDLINE=98195739; PubMed=9534249;
RA   Labo M., Gusberti L., Rossi E.D., Speziale P., Riccardi G.;
RT   "Determination of a 15437 bp nucleotide sequence around the inhA gene
RT   of Mycobacterium avium and similarity analysis of the products of
RT   putative ORFs.";
RL   Microbiology 144:807-814(1998).
DR   EMBL; AF002133; AAC46291.1; -.
SQ   SEQUENCE   278 AA;  30646 MW;  AE985A67[illegible] CRC64;

  Query Match             100.0%;  Score 31;  DB 2;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||||||
Db        164 VQRGGR 169

RESULT 4
Q44416
ID   Q44416      PRELIMINARY;      PRT;   394 AA.
AC   Q44416;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ENDOGLUCANASE.
GN   CELC.
OS   Agrobacterium tumefaciens.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=358;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 95164506; PubMed 7860585;
RA   Matthysse A.G., White S., Lightfoot R.;
RT   "Genes required for cellulose synthesis in Agrobacterium
RT   tumefaciens.";
RL   J. Bacteriol. 177:1069-1075(1995).
DR   EMBL; L38609; AAC41433.1; -.
DR   InterPro; IPR002037; Glyco_hydro_8.
DR   Pfam; PF01270; Glyco_hydro_8; 1.
SQ   SEQUENCE   394 AA;  42458 MW;  [illegible] CRC64;

  Query Match             100.0%;  Score 31;  DB 2;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 85;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||||||
Db        262 VQRGGR 267

RESULT 5
Q98BU6
ID   Q98BU6      PRELIMINARY;      PRT;   424 AA.
AC   Q98BU6;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   AMIN TRANSFERASE, NIFS.
GN   MLL5421.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082990; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003006; BAB51876.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   424 AA;  46939 MW;  921[illegible]26523[illegible]FIB CRC64;

  Query Match             100.0%;  Score 31;  DB 16;  Length 424;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||||||
Db         60 VQRGGR 65

RESULT 6
Q9UWU6
ID   Q9UWU6      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9UWU6;
DT   01-MAY-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-[illegible] (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 14.0 KDA PROTEIN.
GN   ORF C22_010.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=DSM 1617 / P2
RA   Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
RA   Confalonieri F., Curtis B., Duguet M., Erauso G., Faguy D.,
RA   Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
RA   Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA   St Jean A., Van Der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
RA   Ragan M.A., Sensen C.W.;
RT   "Gene content and organization of a 281 kbp contig from the genome of
RT   the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y18930; CAB57758.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   130 AA;  14047 MW;  94B19DA20F333197 CRC64;

  Query Match             96.8%;  Score 30;  DB 1;  Length 130;
  Best Local Similarity   83.3%;  Pred. No. 43;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :||||
Db        109 IQRGGR 114


RESULT 7
O74066
ID   O74066      PRELIMINARY;      PRT;   269 AA.
AC   O74066;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 30.7 KDA PROTEIN.
OS   Cenarchaeum symbiosum.
OC   Archaea; Crenarchaeota; Cenarchaeum.
OX   NCBI_TaxID=46770;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B;
RX   MEDLINE=98422450; PubMed=9748430;
RA   Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA   Swanson R.V.;
RT   "Genomic analysis reveals chromosomal variation in natural populations
RT   of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
RL   J. Bacteriol. 180:5003-5009(1998)
DR   EMBL; AF083072; AAC62709.1; -.
DR   InterPro; IPR003115; ParBc.
DR   Pfam; PF02195; ParBc; 1.
DR   SMART; SM00470; ParBc; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   269 AA;  30689 MW;  AA45216CAD16005B CRC64;

  Query Match             96.8%;  Score 30;  DB 1;  Length 269;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :||||
Db         55 IQRGGR 60


RESULT 8
O74043
ID   O74043      PRELIMINARY;      PRT;   273 AA.
AC   O74043;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 31.1 KDA PROTEIN.
OS   Cenarchaeum symbiosum.
OC   Archaea; Crenarchaeota; Cenarchaeum.
OX   NCBI_TaxID=46770;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A;
RX   MEDLINE=98422450; PubMed=9748430;
RA   Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA   Swanson R.V.;
RT   "Genomic analysis reveals chromosomal variation in natural populations
RT   of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
RL   J. Bacteriol. 180:5003-5009(1998)
DR   EMBL; AF083071; AAC62686.1; -.
DR   InterPro; IPR003115; ParBc.
DR   Pfam; PF02195; ParBc; 1.
DR   SMART; SM00470; ParBc; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   273 AA;  31103 MW;  683C6830180F61BF CRC64;

  Query Match             96.8%;  Score 30;  DB 1;  Length 273;
  Best Local Similarity   83.3%;  Pred. No. 95;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :||||
Db         59 IQRGGR 64


RESULT 9
Q9AIN6
ID   Q9AIN6      PRELIMINARY;      PRT;   273 AA
AC   Q9AIN6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   HYPOTHETICAL 29.5 KDA PROTEIN.
GN   2SC10A7.32.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Scott K.J., Harris D.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000?51; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL583?19; CAC32368.1; -.
DR   InterPro; IPR003018; GAF.
DR   Pfam; PF01590; GAF; 1.
DR   SMART; SM00065; GAF; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   273 AA;  29?23 MW;  ?0?5A6?BD4?0BC7? CRC64;

  Query Match             96.8%;  Score 30;  DB 2;  Length 273;
  Best Local Similarity   83.3%;  Pred. No. 9?;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
                :||||
Db         ?3 LQRGGR ?8


RESULT 10
Q93N8?
ID   Q93N8?      PRELIMINARY;      PRT;    72 AA
```

```
AC   Q93N80;
DT   01-DEC-2001 (TrEMBLrel. 14, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   FERREDOXIN.
OS   Streptomyces lavendulae.
OC   Bacteria; Firmicutes; Actinobacteria, Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1914;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21454913; PubMed=11417274;
RA   Chiu H.T., Hubbard B.K., Shah A.N., Eide J., Fredenburg R.A.,
RA   Walsh C.T., Khosla C.;
RT   "Molecular cloning and sequence analysis of the complestatin
RT   biosynthetic gene cluster.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:8548-8553(2001).
RL   EMBL; AF386507; AAK81933.1; -.
RL   SEQUENCE   72 AA;  7942 MW;  6C2BFCC12F998F96 CRC64;

  Query Match             90.4%;  Score 28;  DB 2;  Length 72;
  Best Local Similarity   83.3%;  Pred. No. 62;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :|||||
Db      66 VERGGR 71


RESULT 11
O26530
ID   O26530      PRELIMINARY;      PRT;   134 AA.
AC   O26530;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   HYPOTHETICAL 14.7 KDA PROTEIN.
GN   MTH430.
OS   Methanothermobacter thermautotrophicus.
OC   Archaea; Euryarchaeota; Methanobacteria, Methanobacteriales;
OC   Methanothermobacter.
OX   NCBI_TaxID=145262.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DELTA H;
RX   MEDLINE=98047514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicaire R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrokovski S., Church G.M.,
RA   Daniels C.J., Mao J.I., Rice P., Nolling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
DR   EMBL; AE000827; AAB84936.1; -.
DR   PIR; D69211; D69211.
SQ   SEQUENCE   134 AA;  14672 MW;  5F32AF0C2F352F28 CRC64;

  Query Match             90.4%;  Score 28;  DB 17;  Length 134;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :|||||
Db     108 LQRGGR 113


RESULT 12
Q9BX98
ID   Q9BX98      PRELIMINARY;      PRT;   141 AA.
AC   Q9BX98;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   UBIQUITIN A-52 RESIDUE RIBOSOMAL PROTEIN FUSION PRODUCT 1
DE   (FRAGMENT).
GN   UBA52.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang H., Jhanji Y., Okamoto I.;
RT   "Human ubiquitin A-52 residue ribosomal protein fusion product 1
RT   (UBA52) in salivary epithelial cells.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF318709; AAK31162.1; -.
DR   HSSP; P22248; 1UBI.
DR   InterPro; IPR000626; Ubiquitin.
DR   Pfam; PF00240; ubiquitin; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00299; UBIQUITIN_1; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
KW   Ribosomal protein.
FT   NON_TER       1       1
SQ   SEQUENCE   141 AA;  16175 MW;  2FE1A9B726571D02 CRC64;

  Query Match             90.3%;  Score 28;  DB 4;  Length 141;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :|||||
Db       4 LQRGGR 9


RESULT 13
Q94JK9
ID   Q94JK9      PRELIMINARY;      PRT;   212 AA.
AC   Q94JK9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   P0434B04.14 PROTEIN.
GN   P0434B04.14.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Rice genome sequencing";
RT   clone:P0434B04.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003541; BAB41994.1; -.
SQ   SEQUENCE   212 AA;  22592 MW;  77181AE4B758D546 CRC64;

  Query Match             90.3%;  Score 28;  DB 16;  Length 212;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :|||||
Db     167 LQRGGR 172


RESULT 14
Q913F2
ID   Q913F2      PRELIMINARY;      PRT;   245 AA.
```

```
AC   Q9I3P2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROBABLE SHORT-CHAIN DEHYDROGENASE
GN   PA1470.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC       (SDR) FAMILY.
DR   EMBL; AE004576; AAG04859.1; -.
DR   HSSP; Q12634; 1YBV.
DR   InterPro; IPR002198; ADH_short.
DR   Pfam; PF00106; adh_short; 1.
DR   PRINTS; PR00080; SDRFAMILY.
KW   Complete proteome; Oxidoreductase.
SQ   SEQUENCE   245 AA;  25365 MW;  342A1CF12C50C714 CRC64;

  Query Match             90.3%;  Score 28;  DB 16;  Length 245;
  Best Local Similarity   83.3%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        129 VERGGR 134


RESULT 15
Q927X9
ID   Q927X9      PRELIMINARY;      PRT;   283 AA.
AC   Q927X9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIN2658 PROTEIN.
GN   LIN2658.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596173; CAC97884.1; -.
DR   InterPro; LIN2658;
KW   Complete proteome.
SQ   SEQUENCE   283 AA;  31447 MW;  7E389DC8A1AE1FB7 CRC64;

  Query Match             90.3%;  Score 28;  DB 19;  Length 283;
  Best Local Similarity   83.3%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        169 IQRGGR 174


Search completed: October 22, 2002, 16:04:33
Job time : 3.0399 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2002, 16:03:01 ; Search time [illegible] Seconds (without alignments) 331.710 Million cell updates/sec

Title: US-09-996-738-8
Perfect score: 31
Sequence: 1 VGRGGR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible] AA1984.DAT:*
6: [illegible] AA1985.DAT:*
7: [illegible]
8: [illegible] AA1987.DAT:*
9: [illegible] AA1988.DAT:*
10: [illegible]
11: [illegible]
12: [illegible]
13: [illegible]
14: [illegible]
15: [illegible]
16: [illegible]
17: [illegible]
18: [illegible]
19: [illegible]
20: [illegible]
21: [illegible]
22: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	100.0	6	22	AAB50043	Human alpha1 integ
2	31	100.0	148	22	AAU19634	Human novel extrac
3	31	100.0	148	22	AAU19794	Human novel extrac
4	31	100.0	206	22	ABG29235	Novel human diagno
5	31	100.0	214	22	AAB50042	Human alpha1 integ
6	31	100.0	378	22	AAU47587	Propionibacterium
7	31	100.0	755	22	AAU57634	Propionibacterium
8	31	100.0	1183	[illegible]	AAY[illegible]725	[illegible]
9	30	96.8	35	17	AAR87641	Synthetic HIV pept
10	30	96.8	269	21	AAY90923	Cenarchaeum symbio
11	30	96.8	273	21	AAY90939	Cenarchaeum symbio
12	28	[illegible]	64	22	ABB16230	Human nervous syst
13	28	90.3	145	9	AAP81759	Sequence encoded b
14	28	90.3	212	20	AAY01198	Polypeptide fragme
15	28	90.3	247	12	AAR15828	Modified III specif
16	28	[illegible]	331	22	ABG18373	Novel human diagno
17	28	90.3	376	19	AAW72189	HSV 2 strain 333
18	28	90.3	376	19	AAW72121	HSV 2 strain 333
19	[illegible]	[illegible]	376	19	AAW72044	HSV 2 strain 333
20	28	90.3	557	22	ABB63394	Drosophila melanog
21	28	90.3	826	22	ABB64648	Drosophila melanog
22	28	[illegible]	880	9	AAP80805	Sequence of env pr
23	27	87.1	54	22	AAU44326	Propionibacterium
24	27	87.1	60	22	AAU56907	Propionibacterium
25	27	87.1	61	22	ABB28011	Human peptide #[illegible]
26	27	87.1	61	22	ABB33184	Peptide #690 encod
27	27	87.1	61	22	ABB18649	Protein #648 encod
28	27	87.1	61	22	AAM53979	Human brain expres
29	27	87.1	61	22	AAM66367	Human bone marrow
30	27	87.1	61	22	AAM14237	Peptide #671 encod
31	27	87.1	61	22	AAM26647	Peptide #684 encod
32	27	87.1	61	22	AAM01968	Peptide #650 encod
33	27	87.1	65	22	AAU48572	Propionibacterium
34	27	87.1	65	22	ABG29620	Novel human diagno
35	27	87.1	67	22	AAU41717	Propionibacterium
36	27	87.1	72	22	AAU42962	Propionibacterium
37	27	[illegible]	73	20	AAY04[illegible]	Mycobacterium spec
38	27	87.1	77	22	AAU07727	Human polypeptide
39	27	87.1	85	22	AAU49459	Propionibacterium
40	27	87.1	91	20	AAY12482	Human 5' EST secre
41	27	87.1	100	22	AAU43082	Propionibacterium
42	27	[illegible]	[illegible]	21	AAB[illegible]7206	Human prostate can
43	27	[illegible]	117	22	AAG[illegible]7152	Human polypeptide
44	27	[illegible]	1[illegible]	2[illegible]	AA[illegible]	Zea mays protein f
45	27	[illegible]	1[illegible]	21	AAB[illegible]	Human prostate can

ALIGNMENTS

RESULT 1
AAB50043
ID AAB50043 standard; peptide; 6 AA.
XX
AC AAB50043;
XX
DT [illegible] 2001 (first entry)
XX
DE Human alpha1 integrin domain epitope.
XX
KW [illegible] antibody;
KW [illegible] condition; asthma;
KW [illegible] gastrointestinal;
KW vascular disease; autoimmune; respiratory distress syndrome;
KW endotoxin shock syndrome; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200073[illegible]-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15004.
XX
PR [illegible]-JUN-1999; 99US-0137038.
PR [illegible]-FEB-2000; 2000US-0184336.
XX
PA (BIOG) BIOGEN INC.
XX
PI [illegible] A, Gotwals P, Lobb R, Koteliansky V;
XX
DR WPI; 2001-061148/07.
XX
PT Use of blocking monoclonal antibody capable of binding to an epitope of

PT VLA-1 for treating inflammatory disorders, in particular arthritis
XX
PS Claim 1; Fig 15; 60pp; English.
XX
CC The present invention relates to the use of an alpha1-I domain
CC blocking antibody capable of binding the ligand binding domain of human alpha1
CC integrin domain (VLA-1). The antibody of the present invention is
CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
CC treatment of pain and headaches, or as an antipyretic, for the treatment
CC of fever, gastrointestinal conditions such as inflammatory bowel
CC diseases, Crohn's disease, gastritis and vascular diseases, migraine
CC headaches, periarteritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's
CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple
CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,
CC allergic rhinitis, respiratory distress syndrome, cerebral ischaemia
CC syndrome and atherosclerosis.
XX
Sequence 6 AA;

Query Match 100.0%; Score 41; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGP 6
||||||
Db 1 VQRGGP 6

RESULT 2
AAU19644
ID AAU19644 standard; Protein; 119 AA.
XX
AC AAU19644;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human novel extracellular matrix protein SEQ ID NO: 294.
XX
KW Human; secreted extracellular matrix protein; immunomodulatory;
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antibacterial; ophthalmic; cytostatic;
KW [illegible] immunostimulant; diagnostic; gene therapy;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; bacterial diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infarction; cerebral infarction;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
OS Homo sapiens.
XX
PN WO200155368-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01318.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198129.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 11-AUG-2000; 2000US-0224538.
PR 11-AUG-2000; 2000US-0224539.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226479.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229443.
PR 01-SEP-2000; 2000US-0229444.
PR 01-SEP-2000; 2000US-0229445.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 13-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 25-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236427.
PR 29-SEP-2000; 2000US-0236467.
PR 29-SEP-2000; 2000US-0236468.
PR 29-SEP-2000; 2000US-0236469.
PR 29-SEP-2000; 2000US-0236476.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 25-OCT-2000; 2000US-0241785.
PR 25-OCT-2000; 2000US-0241786.
PR 25-OCT-2000; 2000US-0241787.
PR 25-OCT-2000; 2000US-0241808.
PR 26-OCT-2000; 2000US-0241809.
PR 27-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246522.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251989.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
DR WPI; 2001-465572/50.
DR N-PSDB; AAS41205.
XX
PT Nucleic acid molecules encoding human secreted extracellular matrix
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers
XX
PS Claim 11; SEQ ID No 284; 577pp; English.
XX
CC The invention relates to isolated nucleic acid molecules encoding
CC novel human secreted extracellular matrix proteins (SPs). The
CC polynucleotides and proteins are used to prevent, treat a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. For example, disorders associated with decreased
CC expression of SPs. The SP polynucleotide or a vector expressing them may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cells own genes and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antigens, which may be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The invention relates to examples of immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), cardiovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation.

Query Match 100.0%; Score 81; DB 22; Length 148;
Best Local Similarity 100.0%; Pred. No. 1e-02
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPGGR 6
:::::
Db 116 VLPGGR 121

FISEQ 3
AAU19794
ID AAU19794 standard; protein; 148 AA.
XX
AC AAU19794;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human novel extracellular matrix protein, Seq ID No 444
XX
KW Human; secreted extracellular matrix protein; immunomodulatory;
KW anti-HIV; anticancer; anti-rheumatoid; antisclerotic; cardiant; vascular;
KW cytogen; fertility; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antiarthritic; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancer; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
OS Homo sapiens
XX
PN WO200155368-A1
XX
PD 02-AUG-2001
XX
PF 17-JAN-2001; 2001WO-US01448.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186450.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205519.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226869.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0234063.
PR 14-SEP-2000; 2000US-0234064.
PR 14-SEP-2000; 2000US-0234065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235464.
PR 27-SEP-2000; 2000US-0235831.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236902.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249241.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (SOMA , G-MAN 'IN-MF SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI, 2002-465512/50.
DR N-PSDB, AAS41365.
XX
PT Novel human secreted extracellular matrix proteins (SPs) -
PT protein, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 11; SEQ ID NO 111; 577pp; English.
XX
CC The invention relates to isolated nucleic acid molecules encoding
CC novel human secreted extracellular matrix proteins (SPs). The
CC polynucleotides and proteins are used to prevent, treat a medical
CC condition in mammals, humans, mice, rabbits, goats, horses, cats, dogs,
CC [illegible] sheep. For example, diseases associated with decreased
CC expression of the SP polynucleotide or protein expression may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cellular messenger and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The disorders include, for example, infectious diseases (e.g. HIV
CC (human immunodeficiency virus) infection, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.

CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), cardiovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation,

Query Match 100.0%; Score 71; DB 22; Length 148;
Best Local Similarity 100.0%; Pred. No. 2e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
||||||
Db 116 VQRGGR 121

RESULT 4
ABG29235
ID ABG29235 standard; Protein; 206 AA.
XX
AC ABG29235;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #29226.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-630422/73
DR N-PSDB; AAS93422.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 59594; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat pathologies involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression. Biological activity
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30477 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 206 AA;

Query Match 100.0%; Score 71; DB 22; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.4e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
||||||
Db 101 VQRGGR 106

RESULT 5
AAB50042
ID AAB50042 standard; protein; 214 AA.
XX
AC AAB50042;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human alpha1 integrin domain.
XX
KW Human; alpha1 integrin domain; alpha1beta1 function blocking antibody;
KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
KW bronchitis; headache; antipyretic; fever; gastrointestinal;
KW vascular disease; autoimmune; respiratory distress syndrome;
KW endotoxin shock syndrome; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200072881-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15004.
XX
PR 01-JUN-1999; 1999US-0137038.
PR 29-FEB-2000; 2000US-0185346.
XX
PA (BIOG-) BIOGEN INC.
XX
PI De Fougerolles A, Gotwals P, Cobb R, Koteliansky V;
XX
DR WPI; 2001-061448/07.
XX
PT Use of blocking monoclonal antibody capable of binding to an epitope of
PT VLA-1 for treating inflammatory disorders, in particular arthritis.
XX
PS Example 11; Fig 1; 69pp; English.
XX
CC The present invention relates to the use of an alpha1beta1 function
CC blocking antibody capable of binding an epitope of human alpha1 integrin
CC domain (VLA-1; see AAB50043). The antibody of the present invention is
CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
CC treatment of pain and headaches, or as an antipyretic for the treatment
CC of fever, gastrointestinal conditions such as inflammatory bowel
CC disease, pneumonitis, restenosis and vascular diseases, migraine
CC headaches, periarteritis nodosa, thyroiditis, aplastic anaemia,
CC Wegner's disease, rheumatic fever, type I diabetes, myasthenia gravis,
CC multiple sclerosis, sarcoidosis, nephrotic syndrome, myocardial
CC ischaemia, fibrosing alveolitis, respiratory distress syndrome, endotoxin
CC shock syndrome and atherosclerosis. The present sequence is the human
CC alpha1 integrin domain.
XX
SQ Sequence 214 AA;

Query Match 100.0%; Score 71; DB 22; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.4e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        1 VQRQGR 6
            ||||||
Db       91 VQRQGR 96


RESULT 6
AAU47587
ID   AAU47587 standard; protein; 378 AA.
XX
AC   AAU47587;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #6464.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skerky YAW; Persing DH; Mitchum D; Ward SS; Bhatia A;
PI   L'maisonneuve J; Zhang Y; Jen S; Carter D;
XX
DR   WPI; 2001-616771/71.
DR   N-PSDB; AAS59629.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Example 1; SEQ ID No 8782; 1969pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system. However it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   378 AA;

  Query Match             100.0%;  Score 31;  DB 22;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 VQRQGR 6
            ||||||
Db       91 VQRQGR 96


RESULT 7
AAU57684
ID   AAU57684 standard; protein; 755 AA.
XX
AC   AAU57684;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #18540.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skerky YAW; Persing DH; Mitchum D; Ward SS; Bhatia A;
PI   L'maisonneuve J; Zhang Y; Jen S; Carter D;
XX
DR   WPI; 2001-616771/71.
DR   N-PSDB; AAS59629.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Example 1; SEQ ID No 18929; 1969pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system. However it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   755 AA;

  Query Match             100.0%;  Score 31;  DB 22;  Length 755;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

Qy 1 VQKGGR 6
       ||||||
Db 303 VQKGGR 308

RESULT 8
AAY07728
ID AAY07728 standard; protein; 1183 AA.
XX
AC AAY07728;
XX
DT 01-JUL-1999 (first entry)
XX
DE Armenian hamster alpha-1 integrin subunit protein.
XX
KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;
KW integrin cell surface receptor; capillary; blood vessel; hamster;
KW alpha-1 subunit; alpha 2 subunit.
XX
OS Cricetulus migratorius.
XX
PN WO9916465-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1997; 97WO-US17485.
XX
PR 30-SEP-1997; 97WO-US17485.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (CLAF/) CLAFFEY K P.
PA (DETM/) DETMAR M.
PA (SENG/) SENGER D R.
XX
PI Claffey KP, Detmar M, Senger DR;
XX
DR WPI; 1999-254938/21.
XX
PT Inhibition of tumor angiogenesis through interaction of vascular
PT endothelial growth factor and integrin cell surface receptors
XX
PS Disclosure; Fig 2A-C; 64pp; English.
XX
CC This invention describes a novel method for the inhibition of tumour
CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and
CC integrin cell surface receptors expressed in vasculature of living
CC subjects. The method inhibits new capillary and new blood vessel
CC formation both within a tumour mass itself as well as in the immediately
CC adjacent blood vasculature surrounding the perimeter of the tumour mass
CC Interaction and dependence upon VEGF to induce specific integrin
CC heterodimers in tumour angiogenesis provides a novel method for
CC inhibiting tumour angiogenesis, and unlike prior art relies on the
CC specific inter relationship of VEGF and integrins, rather than
CC concentrating solely on one specific class of protein.
XX
SQ Sequence 1183 AA;

Query Match 100.0%; Score 33; DB 20; Length 1183;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQKGGR 6
       ||||||
Db 242 VQKGGR 247

RESULT 9
AAR87641
ID AAR87641 standard; peptide; 35 AA.
XX
AC AAR87641;
XX
DT 23-JUL-1996 (first entry)
XX
DE Synthetic HIV peptide CLTB 72.
XX
KW HIV; immunogenic; mineral salt adjuvant; glycolipid analogue; maleimidyl;
KW succinimidyl; 2-pyridyldithio; lipoprotein; glycosylamide; antigen;
KW rubella virus; respiratory syncytial virus; MSA; bordetella pertussis;
KW haemophilus influenzae; streptococcus pneumoniae.
XX
OS Synthetic.
XX
PN WO9634308-A1.
XX
PD 21-DEC-1996.
XX
PF 1[illegible]-JUN-1995; 95WO-CA00359.
XX
PR 1[illegible]-JUN-1994; 94GB-8261194.
XX
PA (CONN) CONNAUGHT LAB LTD.
XX
PI Chong P, James OA, Kandil A, Klein MH;
XX
DR WPI; 1996-049408/05.
XX
PT New adjuvant compsns. for modulating immune responses - comprise
PT mineral salt adjuvant and at least one other adjuvant, partic. a
PT glycolipid analogue
XX
PS Example 1; Page 34; 54pp; English.
XX
CC AAR87639-R87640 and AAR88415-R88418 represent synthetic peptides used as
CC antigens in an immunogenic composition of the invention. The
CC immunogenic composition also contains a novel adjuvant. This adjuvant
CC comprises a mineral salt adjuvant (MSA), and at least one other adjuvant.
CC Another immunogenic composition covered within the scope of the
CC invention comprises an antigen covalently linked to a glycolipid
CC [illegible] MSA. The [illegible] having a
CC reactive functional group, such as maleimidyl, succinimidyl or
CC 2-pyridyldithio. The MSA is preferably aluminium hydroxide, aluminium
CC phosphate, calcium phosphate, zinc hydroxide or calcium hydroxide. The
CC other adjuvant may be a glycolipid analogue, an octadecyl ester of an
CC amino acid or a lipoprotein. The glycolipid analogue is preferably the
CC glycosylamide
CC N-(2-deoxy-2-L-leucylamino-beta-D-glucopyranosyl)-N-octadecyldodecanamide
CC acetate. The immunogenic compositions can be used to induce an immune
CC response to antigens such as antigenic determinants of HIV (such as this
CC sequence), rubella virus, respiratory syncytial virus, Bordetella
CC pertussis, Haemophilus influenzae or Streptococcus pneumoniae. The use
CC of two different adjuvants can provide an enhanced immune response where
CC their effect can be additive or synergistic
XX
SQ Sequence 35 AA;

Query Match 96.8%; Score 30; DB 17; Length 35;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQKGGR 6
        :|||||
Db 2[illegible] IQKGGR [illegible]8

RESULT 10
AAY90923
ID AAY90923 standard; Protein; 269 AA.
XX
AC AAY90923;
XX
DT 30-AUG-2000 (first entry)
XX
DE Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:23
XX

```
KW   Cenarchaeum symbiosum; non-thermophilic crenarchaeote; physiology;
KW   characterisation; archae; therapeutic; industrial; laboratory.
XX
OS   Cenarchaeum symbiosum.
XX
PN   WO200018909-A2
XX
PD   06-APR-2000.
XX
PF   29-SEP-1999;   99WO-US22752
XX
PR   29-SEP-1998;   98US-0102294
XX
PA   (DIVE ) DIVERSA CORP.
XX
PI   Swanson RV, Feldman RA, Schleper C;
XX
DR   WPI; 2000-292149/25.
DR   N-PSDB; AAA55198.
XX
PT   New nucleic acids and proteins isolated from the non-thermophilic
PT   crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
PT   physiology of these archae and in therapeutic, industrial or laboratory
PT   techniques
XX
PS   Claim 26; Page 133; 210pp; English.
XX
CC   AAA55186 to AAA55226 and AAY90914 to AAY90951 represent nucleic acids
CC   and proteins isolated from the non-thermophilic crenarchaeote
CC   Cenarchaeum symbiosum. The nucleic acids and proteins identified in
CC   the present invention are useful in characterising the physiology of
CC   these archae and can be used in therapeutic, industrial or laboratory
CC   techniques. AAA55227 to AAA55260 represent promoter sequences from
CC   Cenarchaeum symbiosum. AAA55261 to AAA55263 represent PCR primers and
CC   probes used in examples from the present invention.
XX
SQ   Sequence   269 AA;

  Query Match             96.8%;  Score 30;  DB 21;  Length 269;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :||||||
Db      55 IQRGGR 60


RESULT 11
AAY90939
ID   AAY90939 standard; Protein; 273 AA.
XX
AC   AAY90939;
XX
DT   30-AUG-2000  (first entry)
XX
DE   Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:56.
XX
KW   Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physiology;
KW   characterisation; archae; therapeutic; industrial; laboratory.
XX
OS   Cenarchaeum symbiosum.
XX
PN   WO200018909 A2.
XX
PD   06-APR-2000
XX
PF   29-SEP-1999;   99WO-US22752.
XX
PR   29-SEP-1998;   98US-0102294.
XX
PA   (DIVE-) DIVERSA CORP
XX
PI   Swanson RV, Feldman RA, Schleper C;
XX
DR   WPI; 2000-292149/25.
DR   N-PSDB; AAA55214.
XX
PT   New nucleic acids and proteins isolated from the non thermophilic
PT   crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
PT   physiology of these archae and in therapeutic, industrial or laboratory
PT   techniques
XX
PS   Claim 26; Page 171; 210pp; English.
XX
CC   AAA55186 to AAA55226 and AAY90914 to AAY90951 represent nucleic acids
CC   and proteins isolated from the non-thermophilic crenarchaeote
CC   Cenarchaeum symbiosum. The nucleic acids and proteins identified in
CC   the present invention are useful in characterising the physiology of
CC   these archae and can be used in therapeutic, industrial or laboratory
CC   techniques. AAA55227 to AAA55260 represent promoter sequences from
CC   Cenarchaeum symbiosum. AAA55261 to AAA55263 represent PCR primers and
CC   probes used in examples from the present invention.
XX
SQ   Sequence   273 AA;

  Query Match             96.8%;  Score 30;  DB 21;  Length 273;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           :||||||
Db      59 IQRGGR 64


RESULT 12
ABB16230
ID   ABB16230 standard; Protein; 64 AA.
XX
AC   ABB16230;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human neuronal system related polypeptide PRT ID NO 4887.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antiallergic; antiasthmatic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200159066-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
PR   31-JAN-2000; 2000US-0178065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   26-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217187.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
```

PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229514.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236379.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 09-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0250719.
PR 05-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Paben SM;
XX
DR WPI; 2001-511565/60.
DR N-PSDB; ABA12556.
XX
PT Nucleic acid encoding 7224 human nervous system antigen polypeptides,
PT useful for prevention, treatment and/or diagnosis of cancer
PT cancers and metastases
XX
PS Claim 11; SEQ ID NO 4887; 170pp; English.
XX
CC The invention relates to novel genes (ABA11084 ABA21544) and proteins
CC (ABB14678 ABB03001) useful for prevention, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial infarction

```
CC   (d) wound healing, (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   64 AA;

  Query Match             90.3%;  Score 28;  DB 22;  Length 64;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db         18 VERGGR 23


RESULT 13
AAP81759
ID   AAP81759 standard; protein; 145 AA.
XX
AC   AAP81759;
XX
DT   15-NOV-1990  (first entry)
XX
DE   Sequence encoded by env gene of SIV-1 (continued from p81758).
XX
KW   AIDS; immunogen; antigen; vaccine; diagnostic.
XX
OS   Simian immunodeficiency virus.
XX
PN   WO8805440-A
XX
PD   28-JUL-1988.
XX
PF   15-JAN-1988;   88WO-FR00025.
XX
PR   15-APR-1987;   87FR-0005398
XX
PA   (INSP ) INST PASTEUR(AL17/)
XX
PI   Alizon M,  Montagnier L,  Guetard D,  Clavel F,  Sonigo P,  Guyader M;
PI   Tiollais P,  Chakrabarti L,  Desrosiers R;
XX
DR   WPI; 1988-220290/31
DR   N-PSDB; AAN80860
XX
PT   New peptide(s) with immunological properties of HIV-2 envelope protein -
PT   having the structure of simian immune deficiency virus proteins,
PT   useful in diagnosis and of vaccine components
XX
PS   Disclosure; Fig 1B; 86pp; French.
XX
CC   New peptides which have immunological properties in common with those of
CC   the peptide skeleton of the envelope protein of HIV 2 and also have a
CC   peptide structure in common with that of SIV-1 glycoprotein are claimed.
CC   Antigenic and immunogenic conjugates contg. the peptides and a kit to
CC   detect HIV-2 in biological fluids are new. The peptides are useful for in
CC   vitro diagnosis of HIV-2 infection and some of them can be used as
CC   components of immunogens and vaccines against HIV. Antibodies raised
CC   against them can be used for treatment of AIDS.
XX
SQ   Sequence   145 AA;

  Query Match             90.3%;  Score 28;  DB 9;  Length 145;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db        122 LQRGGR 127


RESULT 14
AAY01198
ID   AAY01198 standard; Protein; 212 AA.
XX
AC   AAY01198;
XX
DT   18-MAY-1999  (first entry)
XX
DE   Polypeptide fragment encoded by gene 16.
XX
KW   Human; secreted protein; gene therapy; protein therapy; tissue; cancer;
KW   tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;
KW   developmental abnormality; foetal deficiency; Alzheimer's disease;
KW   cognitive disorder; schizophrenia; immunological disorder; mood disorder;
KW   immune deficiency disease; respiratory disorder; arthritis; skeletal;
KW   haematopoietic disorder; neural; osteoporosis; metabolic disorders;
KW   cardiovascular; endocrine; gastrointestinal; asthma; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO9901020-A2.
XX
PD   14-JAN-1999.
XX
PF   30-JUN-1998;   98WO-US13608.
XX
PR   12-SEP-1997;   97US-0058663.
PR   01-JUL-1997;   97US-0051381.
PR   01-JUL-1997;   97US-0051480.
PR   12-SEP-1997;   97US-0058598.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Carter KC,  Endress GA,  Feng P,  Rosen CA,  Ruben SM;
XX
DR   WPI; 1999-105783/09.
DR   N-PSDB; AAX22126.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, immune deficiency diseases or blood
PT   disorders
XX
PS   Disclosure; Page 27; 179pp; English.
XX
CC   The invention relates to nucleic acid sequences (AAX22111 to AAX22134)
CC   encoding human secreted proteins (AAY01135 to AAY01158). The secreted
CC   protein gene sequences are deposited with the ATCC under deposit number
CC   ATCC 209118. Host cells comprising recombinant vectors containing the
CC   nucleic acid sequences are used for the recombinant production of the
CC   secreted proteins. The polynucleotide and amino acid sequences are useful
CC   for are useful for preventing, treating or ameliorating medical
CC   conditions e.g. by protein or gene therapy. Pathological conditions can
CC   be also diagnosed by determining the amount of the new polypeptides in a
CC   sample or by determining the presence of mutations in the new
CC   polynucleotides. Specific uses are described for each of the
CC   polypeptides, based on which tissues they are most highly expressed
CC   in, and include developing products for the diagnosis or treatment of
CC   cancer, tumours, developmental abnormalities and foetal deficiencies,
CC   autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,
CC   schizophrenia, immunological disorders, immune deficiency diseases
CC   (AIDS), mood disorders, respiratory disorders, arthritis, asthma,
CC   haematopoietic disorders, neural disorders, skeletal disorders,
CC   osteoporosis, metabolic disorders, cardiovascular disorders, endocrine
CC   disorders or gastrointestinal disorders. The polypeptides are also useful
CC   for identifying their binding partners. The present sequence represents a
CC   polypeptide fragment encoded by a gene of the invention (see description
CC   line for gene number).
XX
SQ   Sequence   212 AA;

  Query Match             90.3%;  Score 28;  DB 20;  Length 212;
```

```
Best Local Similarity   83.3%;  Pred. No. 5.4e+02;
Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              :|||||
Db         27 LQRGGR 32


RESULT 15
AAR15828
ID   AAR15828 standard; Protein; 247 AA.
XX
AC   AAR15828;
XX
DT   12-MAR-1992  (first entry)
XX
DE   Modified Fl-specific scTCR, scTCR2.
XX
KW   Fluorescein; single chain; T cell receptor; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..111
FT                   /label= variable_beta
FT   Domain          135..246
FT                   /label= variable_alpha
FT   Region          112..134
FT                   /label= linker
FT   Misc-difference 10
FT                   /label= PHE -> SER
FT                   /note= "substitution"
FT   Misc-difference 41
FT                   /label= MET -> LYS
FT                   /note= "substitution"
FT   Misc-difference 111
FT                   /label= LEU -> THR
FT                   /note= "substitution"
FT   Misc-difference 152
FT                   /label= ILE -> ARG
FT                   /note= "substitution"
FT   Misc-difference 146
FT                   /label= ILE -> SER
FT                   /note= "substitution"
XX
PN   WO9118019-A.
XX
PD   28-NOV-1991.
XX
PF   10-MAY-1991;   91WO-US03284.
XX
PR   15-MAY-1990;   90US-0523632.
XX
PA   (SQUI ) SQUIBB E R & SONS INC.
PA   (DANA ) DANA FARBER CANCER INST.
XX
PI   Reinherz EL,  Novotny J,  Smiley ST,  Li P,  Ganju R;
XX
DR   WPI; 1991-365188/50.
DR   N-PSDB; AAQ15508.
XX
PT   New soluble single chain T-cell receptors - useful in suppressing
PT   and diagnosing auto-immune disease by binding specific
PT   antigen-MHC complex and treating infectious diseases
XX
PS   Example 2; Page 42; 64pp; English.
XX
CC   To increase the solubility of the soluble, anti-fluorescein scTCR
CC   as in AAQ15129 and AAQ15507, a number of hydrophobic variable region
CC   framework amino acid residues predicted to be surface exposed were
CC   replaced with more hydrophilic amino acid residues, as indicated in
CC   the features.  The sequence (AAQ15508) encoding this protein was
CC   subcloned into the pSS1 secretion vector, which has the polylinker
CC   replaced with the pectate lyase B (pelB) leader sequence of Erwinia
CC   carotovora (with its RBS) and a new polylinker.
CC   The gene product is useful in a compsn. for suppressing autoimmune
CC   disease when administered to a mammal in a therapeutically effective
CC   amt.  Doses are 0.01-1.0 mg/kg/day, pref. ...
XX
SQ   Sequence   247 AA;

  Query Match             ... %;  Score 28;  DB 121;  Length 247;
  Best Local Similarity   83.3%;  Pred. No. 6.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQRGGR 6
              ||:|||
Db        148 VQKGGR 153


Search completed: October 22, 2002, 16:02:20
Job time : 4.00893 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

SW protein protein search, using sw model

Run on: October 22, 2002, 16:09:02 ; Search Time 7.9027 Seconds
(without alignments)
182.329 Million cell updates/sec

Title: US-09-996-738-8
Perfect score: 31
Sequence: 1 VQRGGR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 241628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 241628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /SIDS2_gcg/data/Issued_Patents_AA_00MB.pep:*
2: /SIDS2_gcg/data/Issued_Patents_AA_01MB.pep:*
3: /SIDS2_gcg/data/Issued_Patents_AA_02MB.pep:*
4: /cgn2_6/ptodata/2/iaa/6R_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCT_2_MB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile_2.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	30	96.8	45	4	US-08-750-624-3	Sequence 3, Appli
2	27	87.1	295	4	US-09-342-081-15	Sequence 15, Appl
3	27	87.1	309	4	US-09-342-081-2	Sequence 2, Appli
4	27	87.1	534	2	US-08-691-818B-8	Sequence 8, Appli
5	27	87.1	1065	3	US-08-630-172-9	Sequence 9, Appli
6	27	87.1	1065	4	US-09-375-115-9	Sequence 9, Appli
7	27	87.1	1170	2	US-08-789-078-2	Sequence 2, Appli
8	27	87.1	1179	2	US-08-752-647-2	Sequence 2, Appli
9	27	87.1	1179	2	US-08-476-062A-42	Sequence 42, Appl
10	27	87.1	1170	5	PCT-US95-01886-2	Sequence 2, Appli
11	27	87.1	1170	5	PCT-US96-01814-42	Sequence 42, Appl
12	27	87.1	5087	4	US-09-144-085-1	Sequence 1, Appli
13	27	87.1	6095	4	US-09-144-085-2	Sequence 2, Appli
14	26	83.9	684	4	US-09-564-805-233	Sequence 233, App
15	26	83.9	779	4	US-09-564-805-235	Sequence 235, App
16	26	83.9	781	1	US-08-280-690-2	Sequence 2, Appli
17	26	83.9	1111	1	US-08-317-450B-15	Sequence 15, Appl
18	26	83.9	1111	4	US-08-800-593-15	Sequence 15, Appl
19	26	83.9	1171	1	US-08-445-135-1	Sequence 1, Appli
20	26	83.9	1193	1	US-08-317-450B-13	Sequence 13, Appl
21	26	83.9	1193	4	US-08-800-593-13	Sequence 13, Appl
22	26	83.9	1311	2	US-08-549-108-4	Sequence 4, Appli
23	26	83.9	1311	3	US-08-656-056-4	Sequence 4, Appli
24	26	83.9	1311	4	US-08-954-668-4	Sequence 4, Appli
25	26	83.9	1311	5	PCT-US95-13233-4	Sequence 4, Appli
26	26	83.9	15283	2	US-08-473-139A-2	Sequence 2, Appli
27	25	80.6	52	2	US-08-726-306A-143	Sequence 143, App
28	25	80.6	126	1	US-08-502-[illegible]-26	Sequence 26, Appl
29	25	80.6	1,0	4	US-09-100-[illegible]-26	Sequence 26, Appl
30	25	80.6	262	2	US-07-857-224B-62	Sequence 62, Appl
31	25	80.6	262	2	US-07-857-224B-63	Sequence 63, Appl
32	25	80.6	294	2	US-08-469-583A-25	Sequence 25, Appl
33	25	80.6	295	2	US-08-701-191A-98	Sequence 98, Appl
34	25	80.6	295	2	US-07-857-224B-71	Sequence 71, Appl
35	25	80.6	294	2	US-08-701-191A-25	Sequence 25, Appl
36	25	80.6	298	1	US-08-261-662-2	Sequence 2, Appli
37	25	80.6	298	5	PCT-US95-03525-2	Sequence 2, Appli
38	25	80.6	341	4	US-09-026-814-2	Sequence 2, Appli
39	25	80.6	564	3	US-08-904-871-1	Sequence 1, Appli
40	25	80.6	576	4	US-09-347-801-19	Sequence 19, Appl
41	25	80.6	580	1	US-08-420-236B-19	Sequence 19, Appl
42	25	80.6	580	4	US-08-793-624-19	Sequence 19, Appl
43	25	80.6	580	5	PCT-US95-10194-19	Sequence 19, Appl
44	25	80.6	593	4	US-09-472-470-44	Sequence 44, Appl
45	25	80.6	609	3	US-08-941-445A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-750-624-3
; Sequence 3, Application US/08750624
; Patent No. 6240971
; GENERAL INFORMATION:
; APPLICANT: Kandil, Ali
; APPLICANT: James, Olive A
; APPLICANT: Chong, Pele
; APPLICANT: Kenan, Michel H
; TITLE OF INVENTION: MINERAL SALT AND ANOTHER IMMUNOSTIMULATING COMPOUND
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,624
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael T
; REGISTRATION NUMBER: 34,223
; REFERENCE/DOCKET NUMBER: 1038 660
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595 1155
; TELEFAX: (416) 595 1163
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-750-624-3

Query Match 96.8%; Score 30; DB 4; Length 45;
Best Local Similarity 83.3%; Pred. No. 9.6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VQRGGR 6
:|||||
Db 25 VQRGGR 30

RESULT 2
US-09-342-084-11
; Sequence 11, Application US/09342084
; Patent No. 6251668
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Transcription Coactivators
; FILE REFERENCE: BB-1169-A
; CURRENT APPLICATION NUMBER: US/09/342,084
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: 60/092,656
; EARLIER FILING DATE: July 13, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-342-084-11

Query Match 87.1%; Score 27; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QRGGR 6
||||| |
Db 19 QRGGR 23

RESULT 3
US-09-342-084-2
; Sequence 2, Application US/09342084
; Patent No. 6251668
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Transcription Coactivators
; FILE REFERENCE: BB-1169-A
; CURRENT APPLICATION NUMBER: US/09/342,084
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: 60/092,656
; EARLIER FILING DATE: July 13, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Zea mays
US-09-342-084-2

Query Match 87.1%; Score 27; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QRGGR 6
||||| |
Db 212 QRGGR 216

RESULT 4
US-08-691-814B-8
; Sequence 8, Application US/08691814B
; Patent No. 5981218
; GENERAL INFORMATION:
; APPLICANT: Rio, Marie-Christine
; APPLICANT: Tomasetto, Catherine
; APPLICANT: Basset, Paul
; APPLICANT: Byrne, Jennifer
; TITLE OF INVENTION: Isolated Nucleic Acid Molecule Coding
; TITLE OF INVENTION: as Breast-Specific Markers and in Breast Cancer Prognosis
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave., NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,814B
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,201
; FILING DATE: 09-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steele, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1483.0090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-691-814B-8

Query Match 87.1%; Score 27; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QRGGR 6
||||| |
Db 51 QRGGR 55

RESULT 5
US-08-630-172-9
; Sequence 9, Application US/08630172
; Patent No. 6060054
; GENERAL INFORMATION:
; APPLICANT: Stuart, Dwo
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR I
; TITLE OF INVENTION: LYMPHOCYTE VETO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,172
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,720
; REFERENCE/DOCKET NUMBER: 2879-36

```
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303) 863-9700
;       TELEFAX:  (303) 863-0223
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1065 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-640-172-9

  Query Match             87.1%;  Score 27;  DB 4;  Length 1065;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRGGR 6
           |||||
Db     465 QRGGR 469


RESULT 6
US-09-375-419-9
; Sequence 9, Application US/09375419
; Patent No. 6264950
;  GENERAL INFORMATION:
;    APPLICANT:  Stuerz, Uwe
;    TITLE OF INVENTION:  NOVEL PROTEIN AND PROTEIN FOR T
;    TITLE OF INVENTION:  LYMPHOCYTE VES
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & Mcintosh
;      STREET:  1700 Lincoln Street, 35th Floo
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/375,419
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/640,172
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,000
;      REFERENCE/DOCKET NUMBER:  2879-36
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1065 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-375-419-9

  Query Match             87.1%;  Score 27;  DB 4;  Length 1065;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRGGR 6
           |||||
Db     465 QRGGR 469


RESULT 7
```

```
US-08-640-172-9
; Sequence 9, Application US/08640172
; Patent No. 5852085
;  GENERAL INFORMATION:
;    APPLICANT:  Benedict, Stephen H.
;    APPLICANT:  Graham, Teresa
;    APPLICANT:  Chare, Marcia
;    APPLICANT:  Tibbetts, Scott
;    TITLE OF INVENTION:  T-CAM 1/LFA-3 PEPTIDES FOR INDUCING
;    TITLE OF INVENTION:  IMMUNE TOLERANCE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOVEY, WILLIAMS, SIMMONS & COLLINS
;      STREET:  1100 Walnut St.
;      CITY:  Kansas City
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  64106
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/640,172
;      FILING DATE:  03-FEB-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/229,514
;      FILING DATE:  07-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, John M.
;      REGISTRATION NUMBER:  26262
;      REFERENCE/DOCKET NUMBER:  22833
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (816)474-9050
;      TELEFAX:  816)474-9057
;      TELEX:  434-363
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1170 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      CELL TYPE:  T cell
;      CELL LINE:  HL-60
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..25
;      OTHER INFORMATION:  /label= signal
;      OTHER INFORMATION:  /note= "Signal sequence"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  42..79
;      OTHER INFORMATION:  /label= Repeat
;      OTHER INFORMATION:  /note= "Repeat I"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  82..132
;      OTHER INFORMATION:  /label= Repeat
;      OTHER INFORMATION:  /note= "Repeat II"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  [illegible]
;      OTHER INFORMATION:  /label= Repeat
;      OTHER INFORMATION:  /note= "Repeat III"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  [illegible]
;      OTHER INFORMATION:  /label= Repeat
```

```
      OTHER INFORMATION: /note= "Repeat IV"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 447..508
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat V"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 509..567
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat VI"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 568..629
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat VII"
    FEATURE:
      NAME/KEY: Domain
      LOCATION: 170..349
      OTHER INFORMATION: /label= IDomain
      OTHER INFORMATION: /note= "I Domain"
    FEATURE:
      NAME/KEY: Domain
      LOCATION: 1089..1112
      OTHER INFORMATION: /label= Trans
      OTHER INFORMATION: /note= "Transmembrane Domain"
    FEATURE:
      NAME/KEY: Domain
      LOCATION: 1113..1170
      OTHER INFORMATION: /label= Cyto
      OTHER INFORMATION: /note= "Cytoplasmic domain"
    PUBLICATION INFORMATION:
      AUTHORS: Pigott,
      AUTHORS: Power,
      TITLE: LFA-1 Amino acid sequence (alpha) (from
      TITLE: PMA-stimulated HL-60 cells)
      JOURNAL: The Adhesion Molecule Facts Book
      PAGES: 94-95
      DATE: 1993
      RELEVANT RESIDUES IN SEQ ID NO: 2 FROM 1 TO 1170
US-08-789 078-2

  Query Match 87.1%; Score 27; DB 2; Length 1170;
  Best Local Similarity 100.0%; Pred. No. 1.1e+13;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QEQGP 6
        |||||
Db 489 QEQGR 493


RESULT 8
US-08-752-633-2
; Sequence 2, Application US/08752633
; Patent No. 5863886
  GENERAL INFORMATION:
    APPLICANT: Benedict, Stephen H.
    APPLICANT: Siahaan, Teruna
    APPLICANT: Chan, Marcia
    APPLICANT: Tibbetts, Scott
    TITLE OF INVENTION: ICAM 1/LFA 1 PEPTIDES FOR INDUCING
    TITLE OF INVENTION: IMMUNE TOLERANCE
    NUMBER OF SEQUENCES: 19
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: BOVEY, WILLIAMS, TIMMONS & KIDD
      STREET: 1101 Walnut St.
      CITY: Kansas City
      STATE: MO
      COUNTRY: USA
      ZIP: 64106
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC Compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/752,633
      FILING DATE:
      CLASSIFICATION: 514
    ATTORNEY/AGENT INFORMATION:
      NAME: Griffin, John M.
      REGISTRATION NUMBER: 29,262
      REFERENCE/DOCKET NUMBER: 22013
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (816)474-9050
      TELEFAX: 816)474-9057
      TELEX: 434-868
  INFORMATION FOR SEQ ID NO: 2:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1170 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
    MOLECULE TYPE: protein
    HYPOTHETICAL: NO
    ORIGINAL SOURCE:
      ORGANISM: Homo sapiens
      CELL TYPE: T cell
      CELL LINE: HL-60
    FEATURE:
      NAME/KEY: Region
      LOCATION: 1..25
      OTHER INFORMATION: /label= signal
      OTHER INFORMATION: /note= "Signal sequence"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 32..79
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat I"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 82..132
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat II"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 339..391
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat III"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 392..446
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat IV"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 447..508
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat V"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 509..567
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat VI"
    FEATURE:
      NAME/KEY: Region
      LOCATION: 568..629
      OTHER INFORMATION: /label= Repeat
      OTHER INFORMATION: /note= "Repeat VII"
    FEATURE:
      NAME/KEY: Domain
      LOCATION: 170..349
      OTHER INFORMATION: /label= IDomain
      OTHER INFORMATION: /note= "I Domain"
    FEATURE:
      NAME/KEY: Domain
      LOCATION: 1089..1112
```

```
;       OTHER INFORMATION:  /label= Trans
;       OTHER INFORMATION:  /note= "Transmembrane Domain"
;     FEATURE:
;       NAME/KEY:  Domain
;       LOCATION:  1113..1170
;       OTHER INFORMATION:  /label= Cyto
;       OTHER INFORMATION:  /note= "Cytoplasmic Domain"
;     PUBLICATION INFORMATION:
;       AUTHORS:  Pigott,
.       AUTHORS:  Power,
;       TITLE:  LFA-1 Amino acid sequence (alpha) (from
.       TITLE:  PMA-stimulated HL-60 cells)
;       JOURNAL:  The Adhesion Molecule Facts Book
;       PAGES:  94-95
;       DATE:  1993
.       RELEVANT RESIDUES IN SEQ ID NO:  2  FROM 1  TO 1170
US-08-792-633-2

  Query Match             87.1%;  Score 27;  DB 2;  Length 1170;
  Best Local Similarity   100.0%;  Prod. No. 1.1e-03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRQTP 6
           ||||| |
Db     489 QRGTR 493


RESULT 9
US-08-476-062A-42
; Sequence 42, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;     APPLICANT:  Arnaout, M. Amin
;     TITLE OF INVENTION:  CONTROLLING CELLULAR CONTACT/INTERACTION BY
;     TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;     NUMBER OF SEQUENCES:  53
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  US
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows95
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/476,062A
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/216,081
;       FILING DATE:  21-MAR-1994
;       APPLICATION NUMBER:  07/647,847
;       FILING DATE:  04-JAN-1991
;       APPLICATION NUMBER:  07/549,842
;       FILING DATE:  18-JUN-1990
;       APPLICATION NUMBER:  07/212,574
;       FILING DATE:  28-JUN-1988
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Freeman, John W.
;       REGISTRATION NUMBER:  29,066
;       REFERENCE/DOCKET NUMBER:  P-792
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-5070
;       TELEFAX:  617/542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  42:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1170 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
US-08-476-062A-42

  Query Match             87.1%;  Score 27;  DB 2;  Length 1170;
  Best Local Similarity   100.0%;  Prod. No. 1.1e-03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRQTP 6
           ||||| |
Db     489 QRGTR 493


RESULT 10
PCT/US97/04886-2
; Sequence 2, Application US/PCT/US97/04886
;  GENERAL INFORMATION:
;     APPLICANT:  Benedict, Stephen H.
;     APPLICANT:  Stoltman, Teroma
;     APPLICANT:  Chan, Marcia
;     APPLICANT:  Diefenbach, Scott
;     TITLE OF INVENTION:  ICAM 1/LFA 1 PEPTIDES FOR INDUCING
;     TITLE OF INVENTION:  IMMUNE TOLERANCE
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  HOVEY, WILLIAMS, TIMMONS & COLLINS
;       STREET:  1100 Walnut St.
;       CITY:  Kansas City
;       STATE:  MO
;       COUNTRY:  USA
;       ZIP:  64106
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US97/04886
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Collins, John M.
;       REGISTRATION NUMBER:  26262
;       REFERENCE/DOCKET NUMBER:  22833
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (816)474-9050
;       TELEFAX:  816)474-9057
;       TELEX:  434-463
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1170 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;       CELL TYPE:  E coli
;       CELL LINE:  HL-60
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1..25
;       OTHER INFORMATION:  /label= signal
;       OTHER INFORMATION:  /note= "Signal Sequence"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  42..79
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat 1"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  86..123
;       OTHER INFORMATION:  /label= Repeat
```

```
;       OTHER INFORMATION:  /note= "Repeat II"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  229..291
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat III"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  392..446
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat IV"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  447..508
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat V"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  509..567
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat VI"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  568..629
;       OTHER INFORMATION:  /label= Repeat
;       OTHER INFORMATION:  /note= "Repeat VII"
;     FEATURE:
;       NAME/KEY:  Domain
;       LOCATION:  170..349
;       OTHER INFORMATION:  /label= IDomain
;       OTHER INFORMATION:  /note= "I-domain"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1089..1112
;       OTHER INFORMATION:  /label= Trans
;       OTHER INFORMATION:  /note= "Transmembrane domain"
;     FEATURE:
;       NAME/KEY:  Domain
;       LOCATION:  1113..1170
;       OTHER INFORMATION:  /label= Cyto
;       OTHER INFORMATION:  /note= "Cytoplasmic domain"
;     PUBLICATION INFORMATION:
;       AUTHORS:  Pigott,
;       AUTHORS:  Power,
;       TITLE:  LFA-1 Amino acid sequence (alphaL) (from
;       TITLE:  PMA-stimulated HL-60 cells)
;       JOURNAL:  The Adhesion Molecule Facts Book
;       PAGES:  94-95
;       DATE:  1993
;       RELEVANT RESIDUES IN SEQ ID NO:  2: FROM 1 TO 1170
PCT-US95-04886-2

  Query Match             87.1%;  Score 27;  DB 5;  Length 1170;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRGGR 6
           |||||
Db     489 QRGGR 493


RESULT 11
PCT-US96-01314-42
; Sequence 42, Application PCT/US9601314
;  GENERAL INFORMATION:
;     APPLICANT:  M. Amin Arnaout
;     TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;     TITLE OF INVENTION:  ANTAGONISTS
;     NUMBER OF SEQUENCES:  78
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PC/AT Compatible
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US96/01314
;       FILING DATE:  30-JAN-96
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/480,167
;       FILING DATE:  30-JAN-95
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  John W. Freeman
;       REGISTRATION NUMBER:  29,066
;       REFERENCE/DOCKET NUMBER:  P-6270C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  42:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1170
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US96-01314-42

  Query Match             87.1%;  Score 27;  DB 5;  Length 1170;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 QRGGR 6
           |||||
Db     489 QRGGR 493


RESULT 12
US-09-144-085-1
; Sequence 1, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
;  APPLICANT: Gustafsson, Claes
;  APPLICANT: Botlach, Mary C.
;  APPLICANT: Ashley, Gary
;  APPLICANT: Julien, Bryan
;  APPLICANT: Ziermann, Rainer
;  TITLE OF INVENTION: DNA AND PROTEIN SEQUENCES AND EXPRESSION SYSTEMS
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: 0202.2002/1.20
;  CURRENT APPLICATION NUMBER: US/09/144,085
;  CURRENT FILING DATE: 1998-08-31
;  PRIOR APPLICATION NUMBER: 60/057,893
;  EARLIER FILING DATE: 1998-01-22
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 5087
;   TYPE: PRT
;   ORGANISM: Sorangium cellulosum
US-09-144-085-1

  Query Match             87.1%;  Score 27;  DB 4;  Length 5087;
  Best Local Similarity   83.3%;  Pred. No. 4.6e-03;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           |:|||||
Db    2840 VRRGGR 2845
```

```
RESULT 13
US-09-144-085-2
; Sequence 2, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
;  APPLICANT: Gustafsson, Claes
;  APPLICANT: Betlach, Mary C.
;  APPLICANT: Ashley, Gary
;  APPLICANT: Julien, Bryan
;  APPLICANT: Ziermann, Rainer
;  TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASE AND ENZYME DNA
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: 30962-20020.20
;  CURRENT APPLICATION NUMBER: US/09/144,085
;  CURRENT FILING DATE: 1998-08-31
;  EARLIER APPLICATION NUMBER: 09/010,809
;  EARLIER FILING DATE: 1998-01-22
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;  LENGTH: 6095
;  TYPE: PRT
;  ORGANISM: Sorangium cellulosum
US-09-144-085-2

  Query Match             87.1%;  Score 27;  DB 4;  Length 6095;
  Best Local Similarity   83.3%;  Pred. No. 5.5e+04;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           |:||||
Db    2954 VRRGGR 2959


RESULT 14
US-09-564-805-233
; Sequence 233, Application US/09564805
; Patent No. 6333404
; GENERAL INFORMATION:
;  APPLICANT:  Iavlovlan, Sean V.
;  APPLICANT:  Tong, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 233
;  LENGTH: 684
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-564-805-233

  Query Match             83.9%;  Score 26;  DB 4;  Length 684;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           | ||||
Db     230 VNRGGR 235


RESULT 15
US-09-564-805-235
; Sequence 235, Application US/09564805
; Patent No. 6333404
; GENERAL INFORMATION:
;  APPLICANT:  Iavlovlan, Sean V.
;  APPLICANT:  Tong, David H.F.
;  APPLICANT:  Simard, Jacques
;  APPLICANT:  Rommens, Johanna M.
;  APPLICANT:  Myriad Genetics, Inc.
;  TITLE OF INVENTION:  Chromosome 17p Linked Prostate Cancer Susceptibility
;  TITLE OF INVENTION:  Gene and a Paralog and Orthologous Genes
;  FILE REFERENCE:  2318-258
;  CURRENT APPLICATION NUMBER: US/09/564,805
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  US 60/107,468
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  09/434,382
;  PRIOR FILING DATE:  1999-11-05
;  NUMBER OF SEQ ID NOS:  240
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 235
;  LENGTH: 774
;  TYPE: PRT
;  ORGANISM: Saccharomyces cerevisiae
US-09-564-805-235

  Query Match             83.9%;  Score 26;  DB 4;  Length 774;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VQRGGR 6
           | ||||
Db     235 VNRGGR 240


Search completed: October 22, 2002, 16:02:34
Job time : 2.80357 secs
```